

AUTONOME MONOKULARE OPTISCHE IDENTIFIKATION VON MONDKRATERN  
AUS UNBEKANNTER KAMERAPOSITION UND -LAGE

Vom Fachbereich Produktionstechnik  
der  
UNIVERSITÄT BREMEN

zur Erlangung des Grades  
Doktor-Ingenieur  
genehmigte

Dissertation  
von  
Dipl.-Math.techn. Bolko Maass

Gutachter: Prof. Dr. rer. nat. Hansjörg Dittus, Universität Bremen  
Prof. Dr.-Ing. Udo Frese, Universität Bremen

Tag der mündlichen Prüfung: 28. Oktober 2016



# Zusammenfassung

Im Rahmen von zukünftigen vollautonomen robotischen Raumfahrtmissionen ist die Navigationsfähigkeit im Bezug auf den Zielhimmelskörper eine unverzichtbare Fähigkeit. Bis dato existiert insbesondere die Fähigkeit nur sehr eingeschränkt, durch kostengünstige und robuste optische Sensoren wie Kameras Positions- und Lagemessungen durchzuführen, die das Raumfahrzeug absolut verorten können im Koordinatensystem des Himmelskörpers.

In der vorliegenden Arbeit wird diese Thematik am Beispiel Mond untersucht. Sie hat die Entwicklung von Methoden zum Gegenstand, die die Bestimmung der Position und Lage der Kamera algorithmisch und ohne äußeren Eingriff aus einem Einzelbild der Mondoberfläche erlauben. Ein solches System befähigt ein robotische Raumfahrzeug schließlich dazu, ohne *a-priori*-Wissen über Position, Lage oder Zeit anhand eines statischen Kataloges von bekannten Mondkratern eine initiale kombinierte Positions- und Lageschätzung zu erzeugen.

Zu diesem Zweck wird ein bildverarbeitendes Verfahren vorgestellt, das ohne weitere Eingaben in einem Bild der Mondoberfläche Einschlagskrater detektieren kann und daraus ihre jeweiligen räumlichen (dreidimensionalen) Eigenschaften in Relation zur aufnehmenden Kamera ableitet. Aus diesen Eigenschaften der noch isolierten Messpunkte im Bild wird anhand eines dafür entwickelten iterativen Interpolationsalgorithmus eine dichte Repräsentation der gesamten beobachteten Mondoberfläche generiert. Diese Oberflächenapproximation erlaubt die Auflösung der bis dahin unbekannten relativen räumlichen Entfernungen der beobachteten Krater untereinander, was eine Korrespondenzbildung zu einem mitgeführten Katalog von Kratern mittels Vergleiches von Konstellationen mehrerer Krater ermöglicht.

Als messendes System wird das entwickelte Verfahren sowohl abschnittsweise bezüglich der Fehlerfortpflanzung in der Verarbeitungskette untersucht, als auch im Rahmen einer Ende-zu-Ende-Testkampagne. Dabei wird einerseits die Leistungsfähigkeit und Sensitivität der Methoden in Abhängigkeit von äußeren Parametern wie Beleuchtungsbedingungen und Blickrichtung auf die Mondoberfläche bewertet, als auch die raumfahrtmissionsplanerischen Aspekte behandelt, im Rahmen einer Identifizierung der Abhängigkeit der Erfolgswahrscheinlichkeit von Parametern wie der Vollständigkeit des bekannten Kraterkataloges. In den durchgeführten Tests werden über synthetisch generierte Eingangsdaten Hypothesen zum Fehlverhalten generiert, die anhand von Labortests im Labor TRON (Testbed for Robotic Optical Navigation) des Instituts für Raumfahrtssysteme des DLR in Bremen verifiziert werden.





# Inhaltsverzeichnis

<b>Abkürzungen</b>	<b>i</b>
<b>Abbildungsverzeichnis</b>	<b>iii</b>
<b>Tabellenverzeichnis</b>	<b>ix</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Referenz-Szenario und Problemstellung . . . . .	3
1.2 Literaturrecherche . . . . .	7
1.2.1 State of the Art . . . . .	8
1.2.2 Abgrenzung und Alleinstellungsmerkmale . . . . .	17
1.3 Gliederung dieser Arbeit . . . . .	18
1.4 Relevante Nebenarbeiten . . . . .	19
<b>2 Bildverarbeitende Verfahren zur Kraterdetektion</b>	<b>21</b>
2.1 Übernommene Verfahrensgrundlagen . . . . .	22
2.1.1 Motivation der Erkennungsheuristik . . . . .	22
2.1.2 Implementierung . . . . .	26
2.2 Erweiterungen im Rahmen der Dissertation . . . . .	28
2.2.1 Begriffsdefinitionen . . . . .	29
2.2.2 Verfahren zur Beleuchtungsschätzung . . . . .	31
2.3 Zwischenfazit . . . . .	38
<b>3 Startwertlose Krateridentifikation</b>	<b>41</b>
3.1 Einleitung und Motivation . . . . .	41
3.2 Lochkameramodell und Kamerakoordinatensysteme . . . . .	42
3.3 Räumliche Rekonstruktion der Kraterländer . . . . .	44
3.3.1 Projektive Kegel der Kraterellipsen . . . . .	45

3.3.2	Kamerafeste Oberflächennormalen der Krater . . . . .	47
3.3.3	Kreisförmige Kegelschnitte . . . . .	50
3.3.4	Mehrdeutigkeit der Lösungen . . . . .	51
3.4	Formale Grundlagen zur Oberflächenapproximation . . . . .	52
3.5	Parametrische Oberflächenmodelle . . . . .	56
3.6	Freie Oberflächenmodelle . . . . .	59
3.6.1	Darstellung der Oberflächenpunktmenge . . . . .	60
3.6.2	Existenz der expliziten Oberflächenfunktion . . . . .	62
3.7	Algorithmus zur Berechnung einer Funktion der Oberfläche . . . . .	71
3.7.1	Einleitung . . . . .	71
3.7.2	Definitionsbereich . . . . .	72
3.7.3	Bivariater Spline-Ansatz . . . . .	74
3.7.4	Triangulation . . . . .	76
3.7.5	BERNSTEIN-Polynome und BÉZIER-Splines . . . . .	78
3.7.6	Interpolationseigenschaft . . . . .	87
3.7.7	Biegeenergie-Funktion . . . . .	90
3.7.8	Quadratisches Minimierungsproblem . . . . .	93
3.7.9	Lösungsverfahren für das Minimierungsproblem . . . . .	96
3.7.10	Zwischenfazit . . . . .	100
3.8	Konfigurationsalgorithmus . . . . .	100
3.8.1	Einleitung . . . . .	100
3.8.2	Iterative energieminimale Konfiguration . . . . .	100
3.8.3	Zwischenfazit . . . . .	108
3.9	Korrespondenzen zum Katalog und Kamerapose . . . . .	109
3.9.1	Einleitung . . . . .	109
3.9.2	Daten des Problems . . . . .	111
3.9.3	Affine Invarianten der Katalogelemente . . . . .	113
3.9.4	$k$ -vector-Ansatz zur Hypothesengenerierung . . . . .	114
3.9.5	Bestimmung einer Kamerapose . . . . .	120
3.9.6	Korrespondenzhypothesen für erweiterte Graphen . . . . .	122
3.9.7	Gesamtalgorithmus . . . . .	126
3.10	Zwischenfazit . . . . .	128
<b>4</b>	<b>Fehleranalyse des Identifikationsverfahrens</b>	<b>129</b>
4.1	Einleitung . . . . .	129
4.2	Modellierung der Verfahrensfehler . . . . .	131
4.2.1	Ausgangsfehlermaße . . . . .	134
4.2.2	Geschlossene Abschätzung . . . . .	137
4.3	Fehlerverhalten der Oberflächenrekonstruktion . . . . .	139

4.3.1	Charakterisierung des Tiefenfehlerverhaltens . . . . .	147
4.3.2	Linienschnitte des Eingangsfehlertraumes . . . . .	160
4.3.3	Vollabdeckung des Eingangsfehlertraumes . . . . .	165
4.3.4	Modellierte Schranke für gemischte Ausgangsfehler . .	170
4.4	Fehlerverhalten der Korrespondenzbildung . . . . .	174
4.4.1	Toleranzschwellen der Korrespondenzbildung . . . . .	176
4.4.2	Test unter Vollabdeckung des Eingangsfehlertraumes .	179
4.5	Zwischenfazit . . . . .	182
<b>5</b>	<b>Tests des Gesamtverfahrens</b>	<b>185</b>
5.1	Einleitung . . . . .	185
5.2	Kamera- und Beleuchtungs-Elevation . . . . .	187
5.2.1	Der Posenraum von Kamera und Beleuchtung . . . . .	188
5.2.2	Kraterkatalog . . . . .	190
5.2.3	Künstliche Bilddaten . . . . .	191
5.2.4	Ergebniskategorisierung . . . . .	194
5.2.5	Ergebnisse . . . . .	195
5.2.6	Auswertung und Fazit . . . . .	201
5.3	Mittlere Akquisitionsperiode . . . . .	204
5.3.1	Bildaustauschperiode . . . . .	205
5.3.2	Testdaten und Testdefinition . . . . .	206
5.3.3	Auswertungsmethodik . . . . .	209
5.3.4	Ergebnisse . . . . .	210
5.3.5	Auswertung und Fazit . . . . .	212
5.4	TRON-Labortests . . . . .	215
5.4.1	Versuchsaufbau und Trajektoriendefinition . . . . .	216
5.4.2	Bilddaten . . . . .	219
5.4.3	Auswertungsmethodik . . . . .	219
5.4.4	Ergebnisse, Auswertung und Fazit . . . . .	221
<b>6</b>	<b>Fazit und Ausblick</b>	<b>229</b>
6.1	Fazit . . . . .	229
6.2	Ausblick . . . . .	232
<b>A</b>	<b>Beispielrechnungen</b>	<b>235</b>
A.1	Beispiel für eine lineare Nadirbestimmung . . . . .	235
A.2	Beispiel einer expliziten Oberflächenrekonstruktion in zwei Di- mensionen . . . . .	236

A.3 Beispiel einer expliziten Spline-Oberflächenrekonstruktion in  
drei Dimensionen . . . . . 240

**Literatur** . . . . . **247**

# Abkürzungen

**ATON** Autonomous Terrain-based Optical Navigation

**DEM** Digital Elevation Model, digitales Höhenmodell

**DLT** Direct Linear Transformation

**DOI** Descent Orbit Injection

**EKF** Extended Kalman Filter

**EP*n*P** Efficient Perspective *n*-Point Solver

**ESA** European Space Agency

**FPGA** Field Programmable Gate Array

**GNC** Guidance Navigation and Control

**GRAIL** Gravity Recovery And Interior Laboratory

**IMU** Inertial Measurement Unit, inertiale Messeinheit

**LADEE** Lunar Atmosphere and Dust Environment Explorer

**LCROSS** Lunar Crater Observation and Sensing Satellite

**LEO** Lunarer Erkundungsorbiter

**LOI** Lunar Orbit Insertion

**LRO** Lunar Reconnaissance Orbiter

**MCMF** Moon-Centered Moon-Fixed

**MSER** Maximally Stable Extremal Regions  
**NEAR** Near Earth Asteroid Rendezvous  
**P $n$ P** Perspektivisches  $n$ -Punkte-Problem  
**PDI** Powered Descent Initiate  
**QCP** Quaternion-based Characteristic Polynomial  
**ROI** Region of Interest  
**SELENE** Selenological and Engineering Explorer  
**SLAM** Simultaneous Localization And Mapping  
**SVD** Singular Value Decomposition  
**TRON** Testbed for Robotic Optical Navigation

# Abbildungsverzeichnis

1.1	Seitlicher Schnitt durch den Mond und Navigationsphasen während eines typischen Mondsüdpol-Landeszenarios . . . . .	4
2.1	Schnitt durch das physische Profil eines typischen Mondkraters unterlegt mit dem aus der eingezeichneten Beleuchtung resultierenden Helligkeitsprofil . . . . .	23
2.2	Fünf charakteristische morphologische Kategorien von lunaren Kratern . . . . .	24
2.3	Morgenaufnahme des Kraters CLAVIUS auf der erdzugewandten Seite des Mondes . . . . .	25
2.4	Charakteristische Hell-Dunkel-Zonen in einem unregelmäßigen Krater der Kategorie ALC mit dem Vektor der lokal dominanten Beleuchtungsrichtung . . . . .	25
2.5	Laufende Schwellwertzerlegung eines Bildausschnittes von CLAVIUS-D mit stabilen zusammenhängenden Gebieten nach dem Kriterium der stetigen Schwerpunktbewegung . . . . .	27
2.6	Zerlegung der CLAVIUS-Bildes in seine stabilen Hell-und-Dunkel-Gebiete . . . . .	28
2.7	Größenabhängige unterschiedliche Anzahlen von dunklen und hellen Gebieten pro Krater und vollständiger Zuordnungsgraph nächster und nächstnächster Partnergebiete . . . . .	32
2.8	Krater CLAVIUS überlagert mit dem Richtungshistogramm der Kanten des Nachbarschaftsgraphen aller stabil dunklen und hellen Gebiete . . . . .	34
2.9	Richtungshistogramm der Kanten des Nachbarschaftsgraphen, Verlauf von $\sigma$ und Minimum der Schätzfunktion der Beleuchtungsrichtung . . . . .	36

3.1	Kamera- und Bildkoordinatensysteme . . . . .	43
3.2	Zusammenhang eines räumlichen Kreises mit einer Bildellipse über den projektiven Kegel . . . . .	45
3.3	Herstellung eines kreisartigen Kegelschnittes durch Rotation einer Ebene über einen Eigenvektor der Kegelquadrik . . . . .	48
3.4	Direkte Parametrisierung des kreisartigen Kegelschnittes durch Linearkombination der Eigenvektoren der Kegelquadrik . . . . .	49
3.5	Mehrere Lösugen für den Normalenvektor der Schnittebene . . . . .	51
3.6	Doppelte Lösung $n$ vs. $n'$ für drei Krater. . . . .	52
3.7	Verschieden konfigurierte und skalierte Oberflächen für drei Krater in 2D. . . . .	55
3.8	Zwei ähnliche Oberflächen . . . . .	55
3.9	Verteilung der zwei ersten Kugelkoordinaten der rekonstruierten Normalen für ein Beispielbild . . . . .	57
3.10	Oberflächendynamik verhindert Nadir-Rekonstruktion . . . . .	60
3.11	Stellen ohne kartesische Umkehrbarkeit . . . . .	62
3.12	Von Bild- nach sphärischen Kamerakordinaten . . . . .	66
3.13	Perspektivische Projektion von räumlichen Kratern in eine Bildebene und Zusammenhang mit sphärischen Koordinaten. . . . .	67
3.14	Darstellung des Rechtecks des Bildrandes nach Transformation in sphärische und virtuelle Bildkoordinaten . . . . .	74
3.15	Die ins Bild transformierte Triangulation über die Zentren der größten erkannten Krater . . . . .	78
3.16	Transformation der Zentren in das virtuelle Bildkoordinaten- system und ihre Triangulation . . . . .	78
3.17	Baryzentrische Koordinaten mit Isolinien auf einem Dreieck . . . . .	80
3.18	Das Kontrollnetz des BERNSTEIN-Polynoms vom Grad 3 . . . . .	82
3.19	Geometrische Stetigkeitsbedingungen eines Splines . . . . .	83
3.20	Geometrische Bedingungen zur stetigen Differenzierbarkeit ei- nes Splines . . . . .	83
3.21	Beispieldiagramm Zuordnungsproblem Kraterkatalog zu rekon- struierten Daten der beobachteten Krater mit Datenfluss. . . . .	110
3.22	Abzählung innerer Kanten von kleinen Graphen und aller Win- kel zwischen ihnen . . . . .	123
4.1	Datenfluss im Navigationskreis . . . . .	130
4.2	Fehlerquellen im Navigationskreis . . . . .	132
4.3	Systematische Generierung und Störung von Testszenarien mit Datenfluss zwischen den Verfahrensstufen . . . . .	142



4.4	Diskretisierung der Störung des Ellipsenzentrums . . . . .	143
4.5	Variation des Mittelwertparameters der Ellipsen-Mittelpunktstörungsverteilung . . . . .	151
4.6	Variation des Varianzparameters der Ellipsen-Mittelpunktstörungsverteilung . . . . .	153
4.7	Variation des Mittelwertparameters der Ellipsen-Halbachsenstörungsverteilung . . . . .	155
4.8	Variation des Varianzparameters der Ellipsen-Halbachsenstörungsverteilung . . . . .	157
4.9	Variation der mittleren Richtungen der Stör-Rotationen der Ellipsen . . . . .	159
4.10	Nichtlineare Modellierung der Mediane der Ausgangsfehlerverteilungen für rotierende Punktverteilung der Eingangsfehlerstreuung der Ellipsenmittelpunkte. . . . .	161
4.11	Nichtlineare Modellierung der Mediane der Ausgangsfehlerverteilungen für wachsende Standardabweichung der Fehler der Ellipsenmittelpunkte . . . . .	162
4.12	Nichtlineare Modellierung der Mediane der Ausgangsfehlerverteilungen für rotierende Punktverteilung der Eingangsfehlerstreuung der Ellipsenhalbachsen. . . . .	163
4.13	Konstante Modellierung der Mediane der Ausgangsfehlerverteilungen für wachsende Standardabweichung der Fehler der Ellipsenhalbachsen. . . . .	164
4.14	Konstante Modellierung der Mediane der Ausgangsfehlerverteilungen für wandernde Schiefe der Rotations-Eingangsfehlerrichtungen. . . . .	165
4.15	Vollabdeckung des Eingangsfehlerverteilungs-Parameterraumes stratifiziert nach Kamera-Elevation und Eingangsfehlerbetrag für die Ellipsenmittelpunkte . . . . .	167
4.16	Vollabdeckung des Eingangsfehlerverteilungs-Parameterraumes stratifiziert nach Kamera-Elevation und Eingangsfehlerbetrag für die Ellipsen-Halbachsen . . . . .	169
4.17	Modellierte Schranke der realisierten Mediane der Ausgangsfehlerverteilungen für gemischte Eingangsfehlerverteilungen aller Störvariablen zu den getesteten Kamera-Elevationen . . . . .	173
4.18	Approximation der Störung der Graphen-Innenwinkel durch Fehler in der Tiefenrekonstruktion an einem Hilfsdreieck und Zusammenhang mit ihren relativen Fehlern . . . . .	176

4.19	Erfolgsrate der Korrespondenzbildung im Vergleich zwischen ungestörtem Ergebnis und für gemischte Vollabdeckung der Eingangsfehlerverteilungsparameter mit maximaler Eingangsfehlermagnitude . . . . .	179
5.1	Geländeausschnitt des Mondes im Höhenmodell in simpel-zylindrischer Projektion mit Kennzeichnung des ungefähren für die Tests verwendeten kreisförmigen Bereiches . . . . .	188
5.2	Kombinierte Variation von Kamera-Elevation, -Azimut, Sonnelevation und -Azimut für einen runden lunaren Geländeausschnitt von 200km Durchmesser . . . . .	189
5.3	Auszug aus dem künstlichen Bildmaterial der systematischen Tests für Variation der Kamerapose . . . . .	192
5.4	Auszug aus dem künstlichen Bildmaterial der systematischen Tests für Variation der Beleuchtungsbedingungen . . . . .	193
5.5	Erfolgsraten systematischer Tests stratifiziert nach den Poseparametern der Kamera . . . . .	196
5.6	Erfolgsraten für volle Variation von Kamera-Azimut und Kamera-Elevation . . . . .	197
5.7	Frequenz falscher Positiver stratifiziert nach den Poseparametern der Kamera . . . . .	198
5.8	Erfolgsraten systematischer Tests stratifiziert nach den Beleuchtungsparametern der globalen Lichtquelle . . . . .	200
5.9	Erfolgsraten für volle Variation der Beleuchtungsparameter . . . . .	201
5.10	Falsch-Positiv-Raten stratifiziert nach den Parametern der Beleuchtung . . . . .	202
5.11	Erfolgsraten für die kombinierte Variation von Kamera- und Beleuchtungs-Elevation . . . . .	203
5.12	Eine Verteilung synthetischer Krater auf planem Testgelände für Tests der mittleren Akquisitionsperiode mit abgeleitetem Katalog in mehreren Dichtestufen . . . . .	207
5.13	Wartezeitenverteilung in Bildern für vollständigen Bildaustausch und verschiedene Dichtestufen des Kataloges mit Exponentialverteilung zum arithmetischen Mittel und Konfidenzintervall zum Erwartungswert . . . . .	211
5.14	Verhalten der Erwartungswerte und 95%-Quantile der Wartezeitenverteilungen in Abhängigkeit von der Katalogdichte . . . . .	212

5.15	Wartezeitenverteilung in Bildern für variablen Bildaustausch und konstanter Dichte des Kataloges mit Exponentialverteilung zum arithmetischen Mittel . . . . .	213
5.16	Darstellung der Mittelwerte der Akquisitionswartezeiten in Abhängigkeit vom Bildaustausch zwischen jeweils aufeinanderfolgenden Messungen . . . . .	214
5.17	Versuchsaufbau für Trajektorientests im Labor TRON am Institut für Raumfahrtssysteme des DLR in Bremen . . . . .	217
5.18	Trajektorie mit Downrange-Sichtfeldgrenzen für Tests zur Bestimmung der mittleren Akquisitionsperiode im TRON-Versuch	218
5.19	Auszug aus dem aufgenommenen Bildmaterial für die TRON-Tests . . . . .	220
5.20	Wartezeitenverteilungen im TRON-Test und Exponentialmodell zum Stichprobenmittelwert. . . . .	221
5.21	TRON-Test: Gemessene Wartezeiten vor den Akquisitionsergebnissen im Vergleich zu ihren Vorhersagen über die variable Katalogdichte mit Betrachtung der relativen Modellfehler. . .	222
5.22	TRON-Test: Gemessene Wartezeiten vor den Akquisitionsergebnissen im Vergleich zu ihren Vorhersagen über die konstante Katalogdichte mit Betrachtung der relativen Modellfehler. . .	224
5.23	Bildweise Wartezeiten bis zum nächsten Akquisitionsergebnis im TRON-Test . . . . .	225
A.1	Beispiel: Lineares Ausgleichsproblem für die Nadirrichtung . .	236
A.2	kartesische Oberflächendaten mit Zentren- und Normalenvektoren	237
A.3	Polare Oberflächendaten . . . . .	237
A.4	Stützwerte der skalierten polaren Ableitung und ihr interpolierendes Polynom. . . . .	239
A.5	Stützstellen und die integrierten Interpolierenden zu neutraler und optimaler Skala. . . . .	239
A.6	Rekonstruierte Oberflächen für zu neutraler und optimaler Skala in kartesische Koordinaten zurücktransformiert. . . . .	239
A.7	Beispielszenario Spline-Oberflächenrekonstruktion räumlich . .	241
A.8	Beispielszenario Spline-Oberflächenrekonstruktion virtuell . .	242
A.9	Ergebnis der Spline-Oberflächenrekonstruktion im Beispiel . .	245
A.10	Ausgeschriebene Interpolationsgleichung . . . . .	246



# Tabellenverzeichnis

1.1	Anforderungen, Voraussetzungen und Beschränkungen von State-of-the-Art-Verfahren zur Navigation durch Kraterdetektion	15
3.1	Zusammenhang der Multiindizes mit den baryzentrischen Kontrollpunkten . . . . .	81
3.2	$\mathcal{C}^0$ - und $\mathcal{C}^1$ -Übergangsbedingungen in Abhängigkeit von den geteilten Seiten der Dreiecke von rechts nach links. . . . .	85
3.3	Explizite $\mathcal{C}^1$ -Übergangsbedingungen mit Skizzen der Anordnung der Kontrollnetz-Indizes . . . . .	86
3.4	Ableitungen der BERNSTEIN-Polynome nach den virtuellen Bildkoordinaten für alle Multiindizes . . . . .	98
3.5	Zweite partielle Ableitungen der BERNSTEIN-Polynome nach den virtuellen Bildkoordinaten für alle Multiindizes . . . . .	99
3.6	Isolinien und Gradienten der rekonstruierten Oberflächen während der iterativen Konfiguration der Daten . . . . .	102
4.1	Fehlerarten und -Ursachen . . . . .	131
4.2	Elemente der Fehlergraphen . . . . .	149
4.3	Modellparameter für Ausgangsfehlermedian-Verteilungen über Kamera-Elevation bei maximalem Eingangsfehlerbetrag . . . .	171
4.4	Ausgleichsgeraden über den Eingangsfehlerbetrag für die relevanten Parameter des Modells der Ausgangsfehlerverteilungs-Mediane . . . . .	173
4.5	Fehlerkategorien der Hypothesenprüfung in Alg. 11 des Korrespondenzbildungsverfahrens . . . . .	175
4.6	Prozentuale Fehlerraten der Zwischenstufen der Korrespondenzbildung in Kategorien nach Tab. 4.5 . . . . .	180

5.1	Erwartete bildweise Erfolgsraten in Abhängigkeit von den äußeren Szenarioparametern Kamera-Elevation und Beleuchtungs-Elevation . . . . .	204
5.2	Mittelwerte der realisierten Wartezeiten für schrumpfende Bildaustauschinkremente mit Vergleich zur Modellvorhersage . . .	215
5.3	Modellgeschätzte Akquisitionswahrscheinlichkeiten zu fester angenommener Katalogdichte und festem angenommenem Akquisitionszeitfenster . . . . .	226

# Kapitel 1

## Einleitung

Sowohl die bemannte als auch die robotische bzw. telekommandierte Exploration des Mondes ruhte größtenteils während der Periode nach Abschluss des bemannten US-Amerikanischen Apollo-Programms und des unbemannten Sowjetischen Luna-Programms Mitte der 1970er bis ins Jahr 2007. Nur eine Handvoll Missionen fand statt während dieser 30 Jahre, und alle waren Orbiter- oder Flyby-Missionen. Im Vergleich hierzu sahen die gut zehn Jahre seitdem eine regelrechte Explosion internationaler Bemühungen, zum Mond zurückzukehren.

Der Japanische Orbiter SELENE (Selenological and Engineering Explorer) im Jahr 2007, die Chinesische Chang'e-Orbiter, -Lander und Roverserie in den Jahren 2007, 2010 und 2013, der Indische Chandrayaan-Orbiter im Jahr 2009 und die US-Amerikanische Serie von Orbitern LRO (Lunar Reconnaissance Orbiter), LCROSS (Lunar Crater Observation and Sensing Satellite), GRAIL (Gravity Recovery And Interior Laboratory) und LADEE (Lunar Atmosphere and Dust Environment Explorer) in den Jahren 2009, 2011 und 2014 flogen überwiegend in Vorbereitung einer erneuten Mondlandung, oder vollbrachten diese sogar bereits, wie im Fall von Chang'e 3.

Es sind insbesondere Schwellenländer wie Russland mit seinen geplanten Missionen Luna-Glob und Luna-Resurs (als letztendliches Ziel wird die Errichtung einer Mondbasis genannt), Indien mit dem Chandraayan-Projekt und China mit seiner Serie von Chang'e-Landern und Rovern, die diese Rückkehr zum Mond vorantreiben. Die USA, die in den Jahren seit 2007 mit ihren oben genannten tatsächlich realisierten Flugmissionen den bislang größten praktischen Beitrag zur Rückkehr zum Mond lieferten, werden voraussichtlich

## 1.0.0 Einleitung

---

in naher Zukunft nicht mehr die Treiber der internationalen Entwicklung auf diesem Gebiet sein. Das US-amerikanische Constellation-Programm, das die bemannte Rückkehr zum Mond bis 2020 zum Ziel hatte, wurde zwischenzeitlich vom US-Kongress zurückgestellt, auch wenn die Entwicklung des Orion-Raumfahrzeugs fortschreitet. Deutsche und Europäische Anstrengungen wie das Projekt LEO (Lunarer Erkundungsorbiter) und Technologieprogramme wie das Aurora- und Next Lunar Lander-Programm der ESA (European Space Agency) wurden aus finanziellen oder politischen Gründen niemals bis zu Reifegraden getrieben, die mit den oben genannten Missionen vergleichbar wären. Sie wurden erheblich verkleinert oder zugunsten internationaler Kooperationen angepasst oder aufgelöst.

Alle noch laufenden Programme setzen zunächst auf robotische Raumfahrzeuge. Während Orbiter generell von der Erde aus zufriedenstellend verfolgt werden können, sind robotische Landefahrzeuge ganz besonders von einer robusten GNC (Guidance Navigation and Control) und einem hohen Grad von Autonomie des Flugrechnersystems abhängig. Insbesondere für Landemissionen, die besondere Ziele von Interesse punktgenau ansteuern wollen, also über sogenannte „Pin-Point Landing“-Fähigkeiten verfügen müssen, ist neben einer robusten Positions- und Lageregelung die akkurate Zustandsschätzung essentiell. Traditionell wird diese in einem Navigationsfilter wie dem KALMAN-Filter vorgenommen, in dem die zeitversetzt und -verzögert einlaufenden fehlerbehafteten Messwerte verschiedener Navigationssensoren fusioniert werden zu einer konsistenten Schätzung der Position, Lage und Geschwindigkeit des Fahrzeugs.

Solche Zustandsschätzer müssen aufgrund ihrer Funktionsweise immer anhand von hochgenau bekannten Zustandsdaten initialisiert werden, etwa durch von der Erde aus gemessener Telemetrie. Ab dem Zeitpunkt des Kontaktverlustes unterliegen die Schätzer allerdings unvermeidlich einer natürlichen Drift der Schätzung, sofern nicht weitere absolute Referenzmessungen eingebracht werden. Für die Lage eines Raumfahrzeuges können Sternsensoren solche Absolutmessungen bereitstellen. Die Position des Fahrzeugs in Relation zum Mond ist mit gängigen Bordverfahren nicht mehr absolut messbar, ihre Schätzung driftet daher unaufhaltsam auf Werte, die mit den Genauigkeiten eines Pin-Point Landings nicht mehr kompatibel sind.

Absolutmessungen der Position mit Bezug auf den Mond lassen sich anhand von mondoberflächenfesten Landmarken, z.B. Kratern, generieren. Wird ein Bild der Mondoberfläche gemacht und kann ermittelt werden, an welcher Stelle im Bild welche bekannten Krater sichtbar sind, so lässt sich mathematisch einfach auf die Position (und Lage) zurückschließen, aus der



das Bild aufgenommen wurde.

Gegenstand dieser Dissertation ist genau diese notwendige Zuordnung von in einem Bild detektierten Kratern zu einem Katalog bekannter mondfest verorteter Krater, in Abwesenheit einer vorherigen guten Zustandsschätzung. Wie sich herausstellen wird, ist diese Fähigkeit äquivalent zur Fähigkeit der Selbstinitialisierung des Navigationssystems und ist daher ein erheblicher Schritt vorwärts in Richtung vollautonomer robotischer Landefahrzeuge, weg von der Abhängigkeit von externer Initialisierung wie der oben besprochenen Initialisierung von der Erde aus.

## 1.1 Referenz-Szenario und Problemstellung

Eine Mondlandemission lässt sich, beginnend im Transferorbit von der Erde zum Mond, in die folgenden charakteristischen Phasen und Ereignisse unterteilen (vgl. Abb. 1.1):

- 1) Transferorbit von der Erde zum Mond, antriebslos
- 2) LOI (Lunar Orbit Insertion), Verringerung der Geschwindigkeit, um stabilen Orbit um den Mond einzunehmen
- 3) Parkorbit um den Mond, antriebslos, kann mehrere komplette Umläufe oder auch nur einen partiellen Umlauf umfassen
- 4) DOI (Descent Orbit Injection), Verringerung der Geschwindigkeit, um elliptischen Abstiegsorbit einzunehmen
- 5) Descent Orbit, HOHMANN-Transfer zum Abstieg auf Vorlandehöhe, antriebslos
- 6) PDI (Powered Descent Initiate), Moment der Triebwerkseinschaltung zum finalen Abstieg
- 7) Powered Descent, aktiver angetriebener Abbau der horizontalen Orbitgeschwindigkeit des Landers bis zum vertikalen Abstieg auf die Landestelle
- 8) Landung

Auch in Abb. 1.1 eingezeichnet ist der Funkschatten, den der Mond erzeugt bezüglich Telekommandierung und Telemetrie von und zur Erde. In diesem Beispielszenario der Südpol-Landung ist zum Zeitpunkt der LOI keine aktive Kommandierung des Raumfahrzeugs möglich. Das ist normalerweise kein Problem, da der Orbit des Fahrzeugs während des Erde-Mond-Transfers hochgenau bekannt war und daher die Unsicherheit seiner bordseitigen Zustandsschätzung zum Manöverzeitpunkt sehr klein sein sollte.

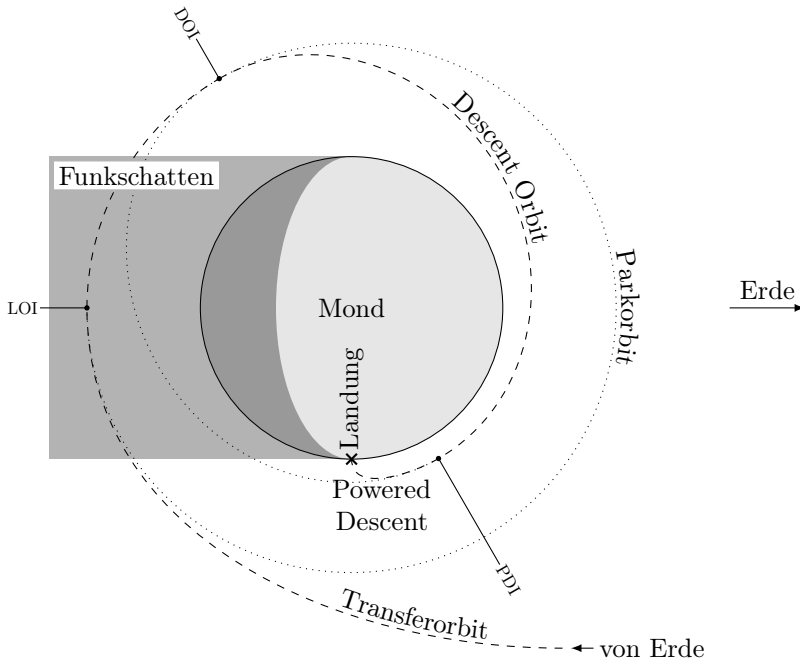


Abbildung 1.1: Seitlicher Schnitt durch den Mond und Navigationsphasen während eines typischen Mondsüdpol-Landeszenarios

Liegt der anvisierte Landeplatz auf anderen Breiten oder an einem Punkt im Funkschatten selbst, diktieren die orbitmechanischen Gesetzmäßigkeiten, dass das Fahrzeug möglicherweise seine Position aus einem sehr lang zurückliegenden bekannten Wert extrapolieren muss. Auch dieses Problem kann üblicherweise gelöst werden, indem über die anfangs bekannte Position und Lage vorhersagbar sichtbare Krater verfolgt werden mittels der Navigationskameras. Die daraus generierten absoluten Referenzmessungen können das Driften des Zustandsschätzers verhindern.

Das bisher nicht gelöste Problem tritt auf, wenn der Zustandsschätzer während einer dieser erdkontaktlosen Phasen erstmals oder neu initialisiert werden muss. Im Kontext der Mondlandung kann eine solche Situation verschiedene Gründe haben:

- Ein unvorhergesehenes Problem ist aufgetreten und das Fahrzeug hat

sich als automatische Sicherheitsmaßnahme zur Sonne ausgerichtet, um den Energiefluss auf die Solarpanele zu maximieren. Währenddessen ist wegen des Manövers und/oder des Sichtverlusts auf die Oberfläche die Navigationsschätzung gedriftet und vorher bekannte Landmarken können nicht wieder aufgenommen werden.

- Die Ausrichtung der Navigationskameras und der Haupttriebwerke zueinander verhindert, dass zu einem der notwendigen Manöver (LOI, DOI, PDI) die Oberfläche sichtbar bleibt, oder die arbeitenden Triebwerke erzeugen zu große Vibrationen, als dass die aufgenommenen Bilder noch verwertbar wären. Nach dem Manöver ist die vorherige Zustandsschätzung mit hoher Wahrscheinlichkeit unbrauchbar.
- Es steht keine Infrastruktur für eine entfernte Positions- und Lageberechnung von der Erde aus zur Verfügung oder keine ist vorgesehen.
- Die Mission umfasst einen Sample Return, d.h. das Fahrzeug muss nach erfolgreicher Landung und den Bodenaktivitäten wieder starten und den Weg zurück zur Erde einschlagen. Geschieht dies auf der Erdausgewandten Seite des Mondes, ist eine Selbstinitialisierung des Zustandsschätzers unabdingbar.

Insbesondere für kommende Kleinmissionen mit entsprechenden Budgetbeschränkungen ist Autonomie in der Poseakquisition ein erheblicher Kostenvorteil und Flexibilitätsfaktor. Auch verfügen viele der derzeit aktivsten oben genannten Schwellenländer schlicht nicht über eine entsprechende technische oder personelle Infrastruktur zur erdgebundenen Verfolgung in der notwendigen Präzision.

Letztendlich ist der Mond natürlich auch nicht das einzige denkbare Ziel für robotische Raumfahrzeuge. Missionen zu entfernten Großasteroiden oder Monden von Mars und Saturn sind bereits im Gespräch oder sogar unterwegs. Auch die jüngeren Kraterformationen des Mars oder Jupiter-Mondes Europa selbst sind trotz atmosphärischer Erosion teilweise geeignet, um die in dieser Dissertation entwickelte Methode unverändert einzusetzen. In solchen Szenarios verbieten die langen Signallaufzeiten von und zur Erde von vornherein eine sichere entfernte Zustandsbestimmung und Steuerung des Fahrzeugs. Grundsätzlich ist die hier entwickelte Methode in keiner Weise an die Parameter des Mondlandeszenarios gebunden. Es ist vielmehr ausreichend, wenn der Zielhimmelskörper geeignete und bekannte Krater auf seiner Oberfläche trägt.

Um die oben motivierte hohe gewünschte Flexibilität der Methode zu ermöglichen, ist noch eine weitere erhebliche Einschränkung der Problemparameter notwendig: Die Berechnung der Position und Lage muss aus einem

### 1.1.0 Referenz-Szenario und Problemstellung

---

Einzelbild der Oberfläche heraus gelingen. Dies hat den einfachen Grund, dass ansonsten Beschränkungen bezüglich sowohl der Kenntnis als auch der Regelung der relativen Ausrichtung der Kamera über eine Bildserie hinweg notwendig wäre. Genau dies erfordert aber bereits initialisierte und funktionierende GNC-Systeme oder ein teilredundantes Zweitsystem und ist daher nicht praktikabel. Weiterhin würde ein solches Mehrbildverfahren eigene Rechenzeitbeschränkungen erzwingen, um Annahmen an die Relativbewegung der Bilder zueinander über die Zeit zu befriedigen.

Das bedeutet natürlich nicht, dass nur ein einziger Versuch gemacht werden darf und dieser mit absoluter Sicherheit zum Ergebnis führen muss. Nur Annahmen über die Relativeigenschaften von mehreren Bildern und Versuchen zueinander sind unzulässig. Die Forderung zu erfüllen, dass jeder Versuch mit an Sicherheit grenzender Wahrscheinlichkeit ein positives Ergebnis liefert, wäre aufgrund der Abhängigkeit von den lokalen Umgebungsbedingungen unmöglich.

Grundsätzlich kann gesagt werden, dass ein Verfahren, das den beschriebenen Ansprüchen genügt, neben der erstmaligen Ermöglichung gewisser Landeszenarien auch als Werkzeug zur Entspannung der Anforderungen an klassische Landemissionen gesehen werden sollte. Da bereits eine Reihe von unbemannten Fahrzeugen erfolgreich auf dem Mond gelandet sind, ist klar, dass durch Abwarten eines günstigen Zeitfensters, durch strikte Planung, umfassende externe Infrastruktur und flexible Anforderungen an die Landeprecision eine solche Landung durchaus bereits möglich ist. Das vorgeschlagene Navigationsverfahren würde aber genau diese Anforderungen entspannen und darüber hinaus für die oben genannten Notfallszenarien einen Sicherheitspuffer darstellen.

Alle diese Anforderungen lassen sich zu einer Kurzformulierung der zentralen Problemstellung dieser Dissertation kondensieren. Diese wird im Laufe der Arbeit immer wieder referenziert und im Kontext des dort aktuellen Standes der Entwicklung präzisiert werden. Um diese Problemstellung vom letztendlichen Navigations- bzw. Posebestimmungsproblem zu differenzieren, das effizienter und genauer direkt im Zustandsschätzer gelöst werden kann, bezieht sich die Formulierung auf die dafür zentrale Aufgabe der Korrespondenzbildung zwischen optischen Detektionen von Kratern und dem mitgeführten Kraterkatalog:

„Aus einem statischen Katalog von Mondkratern und einem einzigen oder wenigen nicht in besonderem zeitlichen Zusammenhang zueinander stehenden geeigneten Bildern der Mondoberfläche, identifiziere ohne weitere Information ausreichend viele der abgebildeten Krater, um auf die Position und Lage der aufnehmenden Kamera rückschließen zu können. Die Anforderungen an die Lage der Kamera zur Oberfläche und an die externe Beleuchtung der Oberfläche sollen dabei minimal sein.“

Die folgende Literaturbesprechung soll dieses Navigationsproblem in den breiteren Zusammenhang mit verwandten Themen aus der Bildverarbeitung und bisher gemachten Anstrengungen in diesem Bereich setzen. Es wird sich zeigen, dass Teilproblematiken der oben genannten Zielstellung schon behandelt wurden, das Gesamtproblem aber nach wie vor nicht befriedigend gelöst ist.

## 1.2 Literaturrecherche

Die optische Detektion von Einschlagskratern ist nicht nur zu Navigationszwecken interessant, denn akkurate Kenntnis über ihre Zahl, Größe und Verteilung liefert wichtige Hinweise zur Entwicklungsgeschichte des abgebildeten Gebietes. Alle solchen Erhebungen mussten in der Frühzeit der Raumfahrt „von Hand“ vorgenommen werden, also von Wissenschaftlern, die auf Abbildungen von Himmelskörpern manuell die Einschlagskrater markierten, zählten und klassifizierten. Die automatische Detektion und Klassifizierung von solchen Einschlagskratern auf verschiedenen Monden im Sonnensystem, dem Mars oder auch Asteroiden erlaubt im Gegensatz dazu den breiten Einsatz von modernen Techniken des Data Minings, um Fragen nach dem Alter oder Ursprung von Geländeformationen oder des ganzen Himmelskörpers zu beantworten.

Tatsächlich ist die aus den Geowissenschaften stammende Literatur und die Anzahl dort entwickelter Methoden zur Kraterdetektion erheblich umfangreicher und auch älter als die aus dem Bereich GNC. Mit den gängigen Verfahren zur Kraterdetektion verwandte Ansätze zur algorithmischen Erkennung von Kreisen in Luftbildmaterial gehen bis in die 80er Jahre zurück (Cross, 1988). Dass solche Methoden erst neuerdings auch zur Navigation eingesetzt werden, hat vielfältige Gründe. Einerseits erreicht die Leistung

### 1.2.1 State of the Art

---

raumfahrttauglicher Rechnersysteme gerade jetzt erst Niveaus, die einen On-board-Einsatz anspruchsvoller algorithmischer Bildverarbeitung in Echtzeit überhaupt möglich macht. Eine moderne strahlungsharte raumfahrttaugliche CPU läuft typischerweise in Taktfrequenzen von unter 100MHz (Cobham Gaisler AB, nodate).

Andererseits haben viele der in der Literatur vorgeschlagenen Methoden sehr starke Voraussetzungen und Einschränkungen bezüglich optischer Parameter des Aufnahmesystems und der Aufnahmebedingungen generell. Hierzu kann man Lage und Blickwinkel der Kamera zählen, ihre Kalibrationsparameter und Eigenschaften des Sensors, Beleuchtungsparameter der Szene, Auswahl der abgebildeten Gebiete, etc. Viele dieser Bedingungen würden eine strikte Missionsplanung und gerade auch eine bereits funktionierende GNC des Fahrzeugs voraussetzen, das die aufnehmende Kamera trägt.

Während diese Verfahren für sich hochinteressant sind, sollen sie hier nicht behandelt werden. Sie implementieren wie gesagt zumeist rechnerisch extrem teure Methoden, wie z.B. Template Matching durch Bildkorrelation, neuronale Netze und Support Vector Machines, Hough-Transformationen, Eigenvektor-Ansätze, etc. Es ist nicht innerhalb eines absehbaren Zeitrahmens damit zu rechnen, dass solche Verfahren auf raumfahrttauglichen Flugrechnern einsetzbar sein werden. Oft sind auch die für ihren Einsatz notwendigen oben genannten Parameterfenster zu einschränkend.

Stattdessen soll daher im folgenden Abschnitt diejenige Klasse von bisher erfundenen Kraterdetektionsverfahren detaillierter besprochen werden, die grundsätzlich für einen solchen Einsatz und daher auch zur Navigation geeignet ist.

### 1.2.1 State of the Art

Auch der Navigationsaspekt der optischen Landmarkendetektion wird seit inzwischen gut fünfzehn Jahren untersucht mit dem Ziel, durch Zuordnung detektierter Landmarken wie Einschlagskratern eine Absolutmessung des Zustandes eines Raumfahrzeugs machen zu können. Zu Beginn der Arbeit am Thema dieser Dissertation im Jahr 2010 existierte ein Großteil der unten besprochenen Forschung noch nicht. Es wird sich aber zeigen, dass auch in den vergangenen fünf Jahren kein Fortschritt in Richtung der voll autonomen Initialisierung des Zustandes erzielt wurde.

Absolut im Sinne der Navigation wird in dieser Arbeit jede Methode genannt, die nicht relativ ist, also nicht nur eine Relativpose liefert mit Bezug auf irgendeinen früheren Zustand, sondern eine absolute Pose mit Bezug

auf einen globalen Referenzrahmen. Dieser Rahmen ist gewöhnlich ein am Zielhimmelskörper befestigtes (also insbesondere kein inertiales) kartesisches Koordinatensystem, da die detektierten Landmarken als Teil der physischen Struktur des Himmelskörpers natürlich nur mit Bezug auf diesen festen Rahmen konstant sind. In manchen der unten zusammengefassten Arbeiten wird diese Art der Navigation „Terrain-relativ“ genannt, meint aber das selbe Prinzip. Viele dieser Arbeiten behandeln auch nicht allein die Kraterdetektion, sondern komplette hybride Navigationssysteme, die eine Form davon als zentrale stützende Positionsmessung integrieren.

Da es eine Reihe von Gruppen von Autoren gibt, die über diesen Zeitraum im Rahmen verschiedener Projekte unterschiedliche Ansätze mit verschiedenen Zielstellungen verfolgten, weisen die dabei entwickelten Methoden eine entsprechend große Vielfalt an verwendeten Algorithmen, realisierter Autonomie und der Menge von benötigtem *a priori*-Wissen auf. Die Liste dieser Verfahren zur Detektion ist lang. Deshalb soll hier nur die Untermenge von Verfahren diskutiert werden, die aufgrund ihrer Eigenschaften für das Zielszenario dieser Arbeit grundsätzlich infrage kommen. Für einen guten Überblick auf dem Stand des Jahres 2009 über im oben genannten Sinne ungeeignete Verfahren sei verwiesen auf (Hanak, 2009, S. 2-7).

Alle unten beschriebenen (kraterdetektierenden) Verfahren bewegen sich in einem von drei Methodenkomplexen der Bildverarbeitung

1. Parameterraumverfahren wie die HOUGH-Transformation
2. Kantendetektionsverfahren wie der CANNY-Detektor
3. Segmentierungsverfahren wie Schwellwertverfahren, Watershed-Methoden oder Intensitäts-Clustering

Der Grund für die Kürze dieser Liste ist darin zu sehen, dass nur diese Methoden aktuell effizient genug implementiert werden können, um den Laufzeitanforderungen auf raumfahrt-tauglicher Rechnerhardware zu genügen. Inklusive dieser Detektionsstufe muss jede Methode zur Kraternavigation im oben genannten Sinne die folgenden drei Kernaufgaben erfüllen:

1. Detektion der Krater in einem aufgenommenen Kamerabild.
2. Zuordnung einer Untermenge dieser Detektionen zu einem mitgeführten statischen global verorteten Katalog von Kratern.
3. Berechnung einer absoluten Pose des Raumfahrzeugs in Bezug auf den Himmelskörper.

## 1.2.1 State of the Art

---

Es wird sich in diesem Abschnitt zeigen, dass jedes dieser Unterprobleme für sich gelöst wurde, sofern gewissen Einschränkungen bezüglich vorhandenem *a priori*-Wissen oder externen Anforderungen an die Regelung des aufnehmenden Fahrzeugs genügt wird. Auch ist es oft möglich, zwei oder alle dieser Schritte in geschlossenen Rückführungskreisen zur gegenseitigen Stützung zu verwenden, und es gibt viele interessante Ansätze hierzu.

Die folgende Recherche fasst die besprochenen Arbeiten nach den Arbeits- oder Forschungsgruppen zusammen, in denen sie entstanden. Sie ist zeitlich aufsteigend geordnet nach der jeweiligen Erstpublikation. Besonderes Augenmerk wird darauf gelegt, zu jeder besprochenen Methode die Voraussetzungen und Beschränkungen in einigen Kernkategorien festzuhalten, die direkt auf den Autonomiegrad der Methode abbilden. Diese sind am Ende dieses Abschnittes ab Seite 15 in Tabelle 1.1 zusammengefasst und motivieren die Abgrenzung dieser Dissertation gegenüber den Methoden aus der Literatur, die im nachfolgenden Abschnitt 1.2.2 formuliert wird.

**University of Southern California, JPL** (Leroy u. a., 2001) — Die erste Veröffentlichung, die ausdrücklich das Thema autonome Absolutnavigation durch Kratererkennung behandelt. Sie bewegt sich speziell im Kontext der Asteroidenlandung. Ihre Methode der Detektion von Kratern basiert auf der Extraktion von Kantenpunkten im Bild (mittels eines CANNY-Kantendetektors) und einem darauf aufbauenden Tensor-Voting, das diese Punkte zu Kantenabschnitten zusammenfasst, die aufgrund ihrer Krümmung plausiblerweise zu einem Kraterrand gehören könnten. Hierein wird dann eine Ellipse eingepasst.

Zur Korrespondenzbildung zwischen den detektierten Kratern und dem bekannten Kraterkatalog wird als bekannt auf dem Asteroiden die Länge und Breite des Bodenpunktes unter dem Raumfahrzeug vorausgesetzt, woraufhin mittels eines mitgeführten 3D-Modells eine Projektion der bekannten Krater erzeugt wird. Zusammen hiermit lässt sich dann über die detektierten Krater eine affine Transformation (Rotation, Translation, Skalierung) im Bildraum ermitteln, die die verbleibenden Freiheitsgrade Flughöhe und Rotation um die optische Achse der Fahrzeugpose auflöst. Die Autonomie der Methode im Sinne dieser Dissertation ist also relativ gering, da alle sechs Freiheitsgrade des Raumfahrzeugs beschränkt sind oder bekannt sein müssen.

**University of Washington, JPL** (Cheng, Johnson u. a., 2003; Cheng und Miller, 2003; Cheng und Ansar, 2005; Ansar und Cheng, 2005) — Die früheste und meistzitierte umfassende Arbeit auf diesem Forschungsgebiet.



Die genannten Publikationen beschreiben die Evolution eines Systems, das viele innovative Heuristiken vereint, die das an sich rechnerisch teure Kraterdetektions- und Zuordnungsproblem stark vereinfachen.

Das zunächst betrachtete Anwendungsszenario ist wiederum eine Asteroïdennavigation, am Beispiel NEAR (Near Earth Asteroid Rendezvous). Ausgegangen wird bei der Bildverarbeitung, wie bei Leroy u. a., von einer Menge aus dem Bild extrahierter Kantenpunkte. Viele dieser Punkte werden dann ausgeschlossen, indem der Bildgradient an dem Kantenpunkt mit der als bekannt vorausgesetzten Beleuchtungsrichtung verglichen wird. Die verbleibenden Kanten werden anhand ihrer Anordnung mit Bezug auf die Beleuchtungsrichtung gruppiert, nach ihrer Konvexität gefiltert und dann für einen Ellipsenfit genutzt. Eine (nichtlinear-) iterative Verfeinerungsstufe passt diese Ellipsen dann genauer auf die Bildgradienten ein. Die Zuordnung zum Katalog bekannter Krater wird dann durch eine HOUGH-artige Abbildung von detektierten und bekannten Krater in den 6-dimensionalen Parameterraum der Kamerapose erreicht (Cheng, Johnson u. a., 2003). Diese Methode ist deswegen rechnerisch zu bewältigen, da das Anwendungsszenario aufgrund der Größe des Zielhimmelskörpers von nur wenigen hundert Kratern ausgeht.

Die Erweiterung dieses Systems zu einem offline-Orbitschätzer wird in (Cheng und Miller, 2003) beschrieben. Hier wird über einen während des Orbits aufgenommenen Bilddatensatz *a posteriori* dessen Orbitparameter geschätzt. Eine wesentliche Neuerung ist der Einsatz konischer Invarianten für das Nachverfolgen von Kratern über mehrere zeitlich abfolgende Bilder hinweg. Es wird die Tatsache ausgenutzt, dass koplanare geschlossene Kurven wie Bildellipsen gewisse affine Invarianten besitzen (Forsyth u. a., 1991). Dieses Verfahren ist aber grundsätzlich nicht wohldefiniert, da nur sehr kleine Differenzposen Abbildungen von Kratern erzeugen, die gut genug durch affine Transformation approximiert werden können.

Zur Anwendung im Rahmen einer Marslandung wird das Verfahren in (Cheng und Ansar, 2005; Ansar und Cheng, 2005) erweitert. Unter der Annahme, dass anstatt von einem sehr unregelmäßig geformten Asteroiden von einem auf den für die Mission relevanten Skalen nahezu planen Zielgelände ausgegangen werden kann, ist hier die Koplanaritätsannahme auch im Dreidimensionalen beinahe gegeben, und nicht nur im Bildbereich. Durch die hierdurch ermöglichte Korrespondenzbildung über die affinen Invarianten und durch zusätzlichen Einsatz eines schnellen Verfahrens der Posebestimmung (Ansar und Daniilidis, 2003) ist für dieses Szenario damit eine beinahe vorwissenfreie Poseinitialisierung möglich. Nur noch Kenntnis der Beleuchtungsrichtung ist erforderlich für die Kraterdetektion, hierfür genügt eine

### 1.2.1 State of the Art

---

sehr grobe Initialpose. Grundsätzlich liegen genügend Daten vor, um diese Gruppierung auch stochastisch vorzunehmen, man kann also bei erfüllten Voraussetzungen von einem echten autonomen System sprechen (Hanak, 2009, S. 9).

**MIT, Draper Laboratory** (Hale, 2007; Singh und Lim, 2008) — Die Master Thesis von Hale aus dem Jahr 2007 ist die erste anzutreffende Arbeit, die ausdrücklich das „Lost-in-Space“-Szenario als Ziel formuliert. Sie beschreibt leider nicht die verwendete Methode zur Detektion der Krater, setzt aber eine nadirblickende Kamerakonfiguration voraus. Ob dies aus Anforderungen der Bildverarbeitung oder aus dem verwendeten Korrespondenzbildungsansatz erwächst, bleibt unklar. Die Methode ist grundsätzlich zur Selbstinitialisierung fähig, solange diese Anforderung und auch die unbekannten Anforderungen der Kraterdetektion erfüllt sind. Als Verfahren zur freien Korrespondenzbildung zum Kraterkatalog wird der  $k$ -Vector Ansatz verwendet (Samaan, Mortari und Junkins, 2006). Eine Publikation von Singh und Lim baut auf dieser Arbeit auf, implementiert aber einen Prädiktor-Korrektor-Ansatz ohne eine Initialisierungskomponente auf Basis eines EKF (Extended Kalman Filter) (Singh und Lim, 2008).

Zur Kraterdetektion wird hier erstmals ein Segmentierungsverfahren eingesetzt. Es ist aber nicht „dynamisch“ im Sinne der unten besprochenen Methode von Vincent Simard Bilodeau oder des vom Autor entwickelten Verfahrens (Maass, 2010). Man muss daher damit rechnen, dass diese Methode nicht robust gegenüber variablen Kontrastbedingungen ist oder sich auf rechnerisch teure Vorprozessierung verlässt. Die Blob-Extraktion wird letztendlich auch nur zur Plausibilitätsbewertung von durch Kantenerkennung gestützten Deskriptorellipsen eingesetzt.

**University of Texas, NASA Johnson** (Hanak, 2009; Hanak, Crain und Bishop, 2010) — Hanak stellt in seiner Dissertation ein Verfahren vor, das bei strikter Orientierung der optischen Achse am Nadir fähig zur Selbstinitialisierung ist, und zwar mondglobal. Dies steht im Gegensatz zur Methode von Ansar und Cheng, die lediglich für einen kleinen Terrainabschnitt im Bereich der bevorzugten Landestelle und bei vorliegendem nahezu vollständigen Katalog zuverlässig funktioniert.

Die Bildverarbeitung zur Kraterdetektion beginnt mit einer Kantenextraktion mittels eines DERICHE-Filters. Die Reorientierung der Bildgradienten an den Kantenpunkten anhand bekannter Beleuchtung im Bild generiert dann geeignete Eingangsdaten für eine zirkuläre HOUGH-Transformation.

Zur autonomen Schätzung des Beleuchtungsazimuts in den Kamerabildern wird ein texturbasiertes Verfahren eingesetzt (Varma und Zisserman, 2004). Dieses Verfahren ist jedoch strikten Voraussetzungen unterworfen, die für typische Navigationsbilder der Mondoberfläche praktisch nie gegeben sind (Hanak, 2009, S. 22). Daher ist dies nur als Fallback-Lösung beschrieben, Hanak bevorzugt eine unabhängige Lageschätzung aus einem Star Tracker und dadurch eine Ermittlung der Lichtrichtung durch Ephemeriden-Lookups nach Bordzeit.

Die durch die HOUGH-Transformation generierten Kandidatenkreise im Bild werden anhand einer Methode iterativ verfeinert, die der von Cheng, Johnson u. a. ähnelt. Es folgen Heuristiken, die über Betrachtungen der Intensitätsverteilung innerhalb der hypothetisierten Kraterbereiche falsche Positive ausschließen sollen. Ein Vergleich der im Bild detektierten Krater mit dem mitgeführten Katalog wird über Kratertripel gemacht, die wieder über einen  $k$ -Vector in linearer Zeit durchsuchbar werden. Trial-and-Error-Erweiterung der Konstellationen auf vier Krater erzeugt dann verlässliche Korrespondenzen.

Da das Verfahren eine sehr genau am Nadir orientierte optische Achse voraussetzt, geht Hanak davon aus, dass mindestens eine weitere externe Messung wie z.B. von einem Horizontsensor vorliegt, die eine solche Orientierung des Raumfahrzeugs zuverlässig ermöglicht. Bis auf diese Einschränkung ist die vorgestellte Methode also fähig zur Initialisierung eines Navigationsfilters.

**Université de Sherbrooke, NGC Aerospace, Thales Alenia, ESA** (Vincent Simard Bilodeau, Clerc, De Lafontaine u. a., 2010; Simard Bilodeau u. a., 2012; Vincent Simard Bilodeau, Clerc, Draï u. a., 2014; Vincent Simard Bilodeau, 2015) — Die im September 2015 veröffentlichte Dissertation von Bilodeau fasst die Tätigkeit dieser Arbeitsgruppe zusammen (Vincent Simard Bilodeau, 2015). Sie beschreibt allerdings wiederum ein komplettes integriertes Navigationssystem, das Daten einer IMU (Inertial Measurement Unit, inertielle Messeinheit) mit optischen absoluten Messungen aus einem Star Tracker, einem Laser-Altimeter und einem Kratererkennungs-System in einem EKF fusioniert. Sie konzentriert sich nicht hauptsächlich auf die Entwicklung der Kraterdetektion oder der Kraterzuordnung. Der diesbezüglich relevante Teil der Arbeit wurde erstmals 2010 publiziert (Spigai, Clerc und Vincent Simard Bilodeau, 2010).

Methodisch ist die dort vorgestellte Kraterdetektion anstatt der von Leroy u. a. oder Cheng, Johnson u. a. verwendeten rein kantenbasierten Verfahren auf einem Segmentierungsansatz aufgebaut, der durch detektierte Kanten im Bild unterstützt wird. Die Idee, ein adaptives Segmentierungsverfahren einzusetzen,

entstand offenbar parallel zur selben Idee aus der Diplomarbeit des Autors (Maass, 2010). Die algorithmische Umsetzung der beiden Segmentierungen unterscheidet sich aber deutlich. Spigai, Clerc und Bilodeau verwenden zur Extraktion ein k-Means Clustering über die Bildintensitäten, während das vom Autor entwickelte Verfahren einer schlankeren Version der MSER (Maximally Stable Extremal Regions) (Matas u. a., 2004) ähnlicher ist. Auch erfordert ihr Verfahren das Zwischenspeichern der vollständigen isolierten ROIs (Regions of Interest) und noch weitere Nachbereitungsschritte.

Vernachlässigt wird in der diskutierten Methode allerdings die Problematik, dass die Beleuchtungsrichtung im Bild zwar zur Paarung extrahierter Segmente benötigt wird, um eine erfolgreiche Detektion eines Kraters zu ermöglichen, diese aber ohne Wissen über die Kamerapose nicht bekannt sein kann. Dies geschieht mit der Begründung, es gäbe keine realistischen Fälle, in denen der Navigationsfilter keine akkurate Poseschätzung liefern könnte oder autonom initialisiert werden müsste (Vincent Simard Bilodeau, 2015, S. 125). Die Methode ist daher ausdrücklich nicht fähig, sich selbst zu initialisieren. Vielmehr verwendet sie einen aus einem EKF rückgeführten Zustand zur projektiven Prädiktion der Abbildungspunkte bekannter Krater und nimmt dann eine Zuordnung im Bildraum vor. Hieraus lässt sich dann separat oder durch enge Kopplung im Filter ein Update der Pose generieren.

Ein Ausfall oder übermäßiges Driften des Navigationsfilters ist daher nicht vorgesehen, dadurch ist der Raum möglicher Missionsparameter erheblich eingeschränkt.

**ONERA, ESA, Astrium, GMV, Harbin IoT** Es existieren weitere Ansätze, die aber aus dem einen oder anderen Grund nicht relevant sind für diese Arbeit. Eine jüngere Publikation aus dem chinesischen Raum beschreibt beispielsweise ein System, das das Verfahren von Ansar und Cheng mit einem Segmentierungsansatz fusioniert (Yu, H. Cui und Tian, 2014). Leider lässt das publizierte Testdesign des Verfahrens keine Rückschlüsse über die tatsächliche Robustheit des Ansatzes gegenüber Blickwinkelelevation zu. (Zur Verifikation einer auf affinen Invarianten basierenden Methode wird Elevation der Kamera simuliert, indem aus Nadirrichtung aufgenommenes Bildmaterial mittels einer Homografie verzerrt wird.)

Auch andere Landmarken als Einschlagskrater werden von anderen Autoren in ähnlicher Weise zur Posebestimmung eingesetzt. Hier seien Entwicklungen auf Basis von Corner Trackern und Prädiktion solcher Features über 3D-Modelle genannt (Van Pham, Lacroix und Devy, 2008; Van Pham, Lacroix, Devy u. a., 2010), bzw. darauf aufbauend (Delaune, Le Besnerais, Sanfourche,

Plyer u. a., 2011; Delaune, Le Besnerais, Sanfourche, Voirin u. a., 2012). Auch diese Methoden bieten keine Fähigkeit zur Selbstinitialisierung, denn sie sind reine Prädiktor-Korrektor-Methoden wie die von Simard Bilodeau und funktionieren nur im Kontext eines konsistenten Zustandsschätzers.

Kombinierte Methoden über nicht-Krater-Landmarken zusammen mit Kratern existieren ebenfalls (Mammarella u. a., 2011). Hier wurde bei GMV in Spanien im Rahmen des ANTARES-Projektes das Zusammenspiel einer solchen Kombination untersucht. Die Methode zur Kraterdetektion ist eine Vereinfachung der Methode von Cheng u.a. und zeigt entsprechende Performance. Das Gesamtsystem benötigt aber wieder eine externe Initialisierung.

Die folgende Tabelle fasst die oben besprochenen Punkte stichwortartig und kategorisiert zusammen, um den direkten Vergleich der Methoden untereinander herauszustellen. Die Einschätzung des jeweiligen Autonomiegrades ist das zentrale Bewertungsmerkmal jeder Methode.

Tabelle 1.1: Anforderungen, Voraussetzungen und Beschränkungen von State-of-the-Art-Verfahren zur Navigation durch Kraterdetektion

Autoren	Erforderliches <i>a priori</i> -Wissen	Methoden	Einordnung, Autonomiegrad
Leroy und Johnson	<ul style="list-style-type: none"> <li>- <b>Position</b> Länge und Breite im Referenzrahmen des Asteroiden</li> <li>- <b>Lage</b> Nadir Pointing</li> <li>- <b>Kamera</b> Lange Brennweite oder großer Abstand zum Asteroiden zur Minimierung nicht-affiner Effekte</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Bildverarbeitung</b> CANNY-Kantendetektor, Tensor-Voting zur Einschränkung auf konvexe Kanten, iterativer Ellipsenfit</li> <li>- <b>Korrespondenzbildung</b> 2D-affine Transformation im Bild zwischen Paaren aus Detektionen und Projektion eines 3D-Modells</li> </ul>	<p><b>Geringe Autonomie</b> Wegen großer Einschränkungen bzgl. Pose nicht zur Selbstinitialisierung geeignet</p>

Tabelle wird fortgesetzt auf folgender Seite

Fortsetzung Tabelle 1.1

Autoren	Erforderliches <i>a priori</i> -Wissen	Methoden	Einordnung, Autonomiegrad
Cheng, Ansar, Johnson u.a.	<ul style="list-style-type: none"> <li>- <b>Gelände</b> Nahezu perfekt planares Gelände</li> <li>- <b>Beleuchtung</b> Nicht spezifiziert, Elevation muss aber geeignet sein, um harte Schatten zu erzeugen</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Bildverarbeitung</b> Kantenextraktion durch CANNY-Detektor, Kantengruppierung nach Beleuchtungsrichtung und Intensitätsverlauf, direkter Ellipsenfit, iterative Ellipsenverfeinerung</li> <li>- <b>Korrespondenzbildung</b> Projektive Prädiktion und Korrektur (2003, nicht autonom) oder über projektive konische Invarianten (2005, initialisierungsfähig)</li> </ul>	<b>Hohe Autonomie</b> Unter den Voraussetzungen an das Gelände. Kenntnis der Beleuchtungsrichtung nützlich aber nicht strikt notwendig, selbst-initialisierungsfähig
Hale, Sing, Lim	<ul style="list-style-type: none"> <li>- <b>Position, Lage</b> Sehr grobe Kenntnis, um Beleuchtungsrichtung im Bild vorherzusagen, Nadir pointing</li> <li>- <b>Beleuchtung</b> Grobe Kenntnis der Richtung im Bild</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Bildverarbeitung</b> Blob-Extraktion nach Schwellwertbildung für feste Schwellwerte, Ellipsenfit in Kanten aus Sobel-Filter</li> <li>- <b>Korrespondenzbildung</b> <i>k</i>-Vector-Ansatz, Bildraum-Matching nach Rückprojektion</li> </ul>	<b>Teilweise Autonomie</b> Anforderung Nadir Pointing und Beleuchtungsrichtung macht echte Selbstinitialisierung unmöglich, möglicherweise lösbar über Trial-and-Error
Hanak, Crain, Bishop	<ul style="list-style-type: none"> <li>- <b>Lage</b> Nadir Pointing</li> <li>- <b>Beleuchtung</b> Zwischen 5° und 50° Elevation</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Bildverarbeitung</b> Kantenextraktion über DERICHE-Detektor, Zirkuläre Hough-Transformation, iterative Verfeinerung nach Bildgradienten</li> <li>- <b>Korrespondenzbildung</b> <i>k</i>-Vector-Ansatz mit Ausschlusstests über offline hergeleitete Wahrscheinlichkeitsschwellen</li> </ul>	<b>Hohe Autonomie</b> Vollständig selbst-initialisierungsfähig, sofern Nadir Pointing sichergestellt ist

Tabelle wird fortgesetzt auf folgender Seite

Fortsetzung Tabelle 1.1

Autoren	Erforderliches <i>a priori</i> -Wissen	Methoden	Einordnung, Autonomiegrad
Bilodeau, Clerc, de Lafontaine u.a.	<ul style="list-style-type: none"> <li>- <b>Position</b> Kenntnis bis auf <math>\approx 10^4</math>m erforderlich</li> <li>- <b>Lage</b> Nadir pointing bis <math>45^\circ</math> Elevation bezüglich optischer Achse, Robustheit gegen Fehler nicht angegeben, aber eigentlich wichtiger als bei Position</li> <li>- <b>Beleuchtung</b> Elevation zwischen <math>5^\circ</math> und <math>70^\circ</math> bezüglich Gelände</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Bildverarbeitung</b> Vorbehandlung durch Glättung und Histogrammausgleich, k-Means-Intensitätsclustering mit Distance Transform und Watershed Transform extrahiert Hell-Dunkel-Gebiete, Gruppierung zu Kraterkandidaten über Optimierungsmethode</li> <li>- <b>Korrespondenzbildung</b> Projektive Prädiktion über bekannte Pose, stochastische Zuordnung im Bildraum</li> </ul>	<p><b>Geringe Autonomie</b> Reine Prädiktor-Korrektor-Methode, <i>a priori</i>-Wissen in jeder Stufe erforderlich, keinerlei Selbstinitialisierungsfähigkeit</p>

## 1.2.2 Abgrenzung und Alleinstellungsmerkmale

Die Autonomie der oben besprochenen Verfahren ist vollständig bestimmt durch die folgenden Dimensionen:

- Abhängigkeit von dynamischer Information wie dem zeitaktuellen Zustand des Raumfahrzeugs
- Empfindlichkeit gegenüber dynamischen äußeren Bedingungen wie der Beleuchtung
- Stärke der Annahmen an die Art und Variabilität des beobachteten Geländes

Nach den oben gemachten Recherchen ist die Methode von Hanak hier am relevantesten. Sie erfüllt aufgrund der Abhängigkeit von einer funktionierenden GNC für das Nadir-Pointing jedoch nicht vollständig die gestellten Anforderungen. Die zweite nahekommende Lösung, die von Ansar und Cheng, ist aufgrund der gewählten Methode zur Korrespondenzbildung zwischen Kraterkatalog und Detektionen nicht geeignet, auf einen mondglobalen Kontext erweitert zu werden. Auch sind die gemachten Annahmen bezüglich der Planarität des Geländes in den seltensten Fällen erfüllt.

Eine Methode, die nicht von Zustandsdaten des Fahrzeugs jeglicher Art abhängig ist und minimale Anforderungen an das Pointing stellt, die eine

maximale Unempfindlichkeit gegenüber der Beleuchtung bietet und auf allen Skalen und in einer maximalen Bandbreite anzutreffender Geländeformen funktioniert, stellt also einen echten Fortschritt im Stand der Forschung dar. Soweit dem Autor bekannt, ist die in dieser Dissertation entwickelte Methode in Anspruch und Ausführung tatsächlich einzigartig in ihrer Autonomie, bezogen auf die oben genannten Kriterien.

Über die in dieser Form alleinstellende Fähigkeit zur Selbstinitialisierung des Zustandsschätzers hinaus decken die in dieser Arbeit entwickelten Methoden zusätzlich noch die bereits anderweitig (z.B. von Vincent Simard Bilodeau, Clerc, Drai u. a.) entwickelten Fähigkeiten der Kraterdetektion zum Pose-Update in einem geschlossenen Schätzkreis ab. Ein System, das das neue Verfahren umfasst, müsste also kein weiteres optisches Prädiktor-Korrektor-Verfahren beinhalten, denn alle notwendigen Messungen und Berechnungen werden ohnehin gemacht. Die detektierten Krater könnten in einem klassischen Feature Tracker Verwendung finden oder zur visuellen Odometrie verwendet werden. Hier bietet sich eine Chance zur erheblichen Reduktion von Umfang und Komplexität des im Bordrechner laufenden Programmcodes.

## 1.3 Gliederung dieser Arbeit

Der verbleibende Teil der vorliegenden Dissertation gliedert sich im Wesentlichen in drei Hauptteile: Zunächst werden im Kapitel 2 die bildverarbeitenden Methoden motiviert und erläutert und die Daten definiert, die sie aus den Eingabebildern erzeugen. Das Kapitel 3 behandelt im Detail das Prinzip der Modellschätzung, die Motivation und Wohldefiniertheit der Modelle, die Art und Weise ihrer Konstruktion aus den Daten und ihre Lösung, und schließlich die Methode der Zuordnung von Messungen von Kratern zum mitgeführten Kraterkatalog anhand dieser Modelle.

Das Verhalten der Modellschätzung unter Fehlereinfluss der Bildverarbeitung wird im Kapitel 4 untersucht. Es wird herausgestellt, welche Fehlerarten Modellfehler welcher Größenordnung erzeugen. Dieses Wissen wird genutzt, um die später ermittelten Gesamtausfallraten des Verfahrens zu bewerten. Um die Behauptung zu qualifizieren, das konstruierte Verfahren wäre selbst-initialisierungsfähig im oben besprochenen Sinne, wird im Kapitel 5 zunächst eine systematische Variation der relevanten Szenarioparameter vorgenommen, um eine instantane Erfolgswahrscheinlichkeit abschätzen zu können. Die für reale Missionsplanung wichtige Größe der mittleren erwarteten Zeit bis zur erfolgreichen Initialisierung wird in einer weiteren synthetischen Testreihe



ermittelt.

Als Bestätigung der anhand von Simulation und synthetischem Bildmaterial ermittelten Performancegrößen wird ein repräsentatives Mondlandeszenario im Labor TRON nachgestellt, wo mit einer echten Kamera auch die optischen Effekte auf die Bildverarbeitung wirken.

## 1.4 Relevante Nebenarbeiten

Zur Entwicklung und zum Test der in dieser Arbeit behandelten Methoden war eine Vielzahl von Tätigkeiten notwendig oder hilfreich, die in den folgenden theoretischen Kapiteln aus Platzgründen nicht ausdrücklich behandelt werden.

Hierunter fallen zum einen alle Arbeiten an der Referenzimplementierung des entwickelten Gesamtverfahrens, das zum Zeitpunkt der Niederschrift der Testergebnisse gut 10.000 Programmzeilen ANSI-C im Hauptzweig der Software und jeweils ca. 1500 Zeilen Python- und Matlab-Code in den Test- und Verifikationsgerüsten umfasst.

Darüber hinaus wurde im Rahmen des DLR-internen Projektes ATON (Autonomous Terrain-based Optical Navigation) und des Aufbaus des Labores TRON im DLR-Institut für Raumfahrtsysteme in Bremen umfangreiche Software erstellt zur Sichtung, Interpretation, Aufbereitung und Herstellung von Mondgeländemodellen, die im Laufe der Entwicklung der Navigationsmethoden zur Verifikation genutzt wurden und im Kapitel 5 einen wichtigen Teil der Datenbasis zur Performanceanalyse beitragen. Auf Basis eines Teils dieser Modelle wurde auch eine gern genutzte Demo-Applikation zusammengestellt, bestehend aus einem simulierten Raumfahrzeug in Form einer Handkamera, eines Sonnensimulators in Form einer Baustellenlampe und einer Miniaturfräse eines realen Mondgeländes im Maßstab 1:1.000.000.

Zur Funktion des entwickelten Verfahrens ist ein statischer Kraterkatalog notwendig. Da auf den im Navigationsfall abgebildeten Skalen kein solcher Katalog von zufriedenstellender Dichte frei verfügbar ist, wurde ein eigener erstellt. Hierzu wurde eine weitere Software geschrieben, mittels derer ein Mond-DEM (Digital Elevation Model, digitales Höhenmodell) und simulierte Kamerabilder verwendet werden können, um manuell Krater zu identifizieren und mondglob in einem Katalog abzuspeichern.

Ebenfalls im Rahmen des Projekts ATON fand im Jahr 2015 ein erster Helikopterflugtest statt, währenddessen ein integriertes Rechner- und Sensorpaket getestet wurde, das ebenfalls die in dieser Dissertation entwickelte Methode umfasste. Hierfür wurden Testaufbauten entworfen und es wurde die

#### 1.4.0 Relevante Nebenarbeiten

---

Funktionsfähigkeit des Verfahrens in einem fliegenden Systems nachgewiesen.

## Kapitel 2

# Bildverarbeitende Verfahren zur Kraterdetektion

Die in dieser Arbeit verwendete Methode der bildverarbeitenden Erkennung der Krater wurde im Rahmen der Diplomarbeit des Autors in den Jahren 2009 und 2010 entwickelt (Maass, 2010). Eine erste Implementierung des Verfahrens lieferte brauchbare Ergebnisse, erforderte aber noch Kenntnis wichtiger Parameter, wie der dominanten Beleuchtungsrichtung im Bild, die im Zielszenario dieser Dissertation nicht vorliegen („Lost-in-Space“). Sie umfasste zudem einige rechnerisch sehr aufwändige Schritte zum Ausschluss irrtümlicher Erkennungen ohne realistische Chance zeitnaher Implementierung auf Raumfahrt-Hardware.

Im Rahmen dieser Dissertation wurden die diese überkonservativen Ausschluss-Stufen entfernt bzw. durch die in diesem Kapitel beschriebenen Konsistenzkriterien bzgl. der Beleuchtungsrichtung ersetzt. Eine Publikation aus dem Jahr 2011 stellte einen Teil dieser Neuerungen bereits vor (Maass, Krüger und Theil, 2011). Dieses Kapitel soll in Abschnitt 2.1 kurz den Überblick schaffen über das Kratererkennungsverfahren auf seinem momentanen Stand, ohne Diplomarbeit und die genannte Publikation im Detail zu reproduzieren. Es werden in Abschnitt 2.2 dann nur die danach noch hinzugekommenen Änderungen in Gänze hergeleitet, wie die Auflösung der bestehenden Abhängigkeiten von unmöglich zu kennenden Parametern.

Anstatt anhand konkreter Algorithmen wird die Lösung des Kraterdetektionsproblems deklarativ beschrieben. Das geschieht aus dem Grund, dass die tatsächlich finale Implementierung für einen raumfahrttauglichen Rechner

fast sicher die Verwendung eines FPGA (Field Programmable Gate Array) für diese Lösung umfassen wird. Für einen solchen Prozessor sind nicht unbedingt die selben Algorithmen geeignet, die in der für diese Arbeit erstellte Referenzimplementierung geschrieben wurden, auch wenn letztendlich das selbe Problem gelöst wird.

## 2.1 Übernommene Verfahrensgrundlagen

### 2.1.1 Motivation der Erkennungsheuristik

Wie im Abschnitt 1.2.2 besprochen wurde, hat das in dieser Arbeit verwendete Verfahren zur Kratererkennung den Anspruch, über einen möglichst breiten Bereich von äußeren Bedingungen nutzbare Lösungen zu liefern. Darunter versteht sich die Gesamtheit aller Umwelt- und Missionsparameter, etwa Beschaffenheit des abgebildeten Geländes, Beleuchtungsbedingungen, Lage und Position des Raumfahrzeuges, Abbildungseigenschaften der verwendeten Kamera, etc. Zusätzlich zur maximalen Unempfindlichkeit gegenüber diesen äußeren Parametern soll jetzt noch gesichert bleiben, dass für unterschiedliche Einsatzszenarien nur eine minimale Anzahl innerer Parameter des Algorithmus angepasst werden muss.

Im Rahmen des Entwicklungsprozesses des Kratererkennungsverfahrens wurde dazu schließlich eine Heuristik als sehr robust gegenüber den oben genannten Bedingungen identifiziert, die hier kurz vorgestellt werden soll. Die ihr zugrundeliegende Beobachtung war, dass ein Krater als rundliche Vertiefung im umliegenden Gelände durch das einfallende und aus dem Inneren reflektierte Sonnenlicht auf Kamerabildern stets als ein Gebiet mit einer charakteristischen Konfiguration einer helleren und einer dunkleren Fläche abgebildet wird. Abb. 2.1 illustriert, wie dieses charakteristische Helligkeitsprofil zustandekommt.

Das abgebildete Geländeprofil ist für die große Mehrheit aller anzutreffenden Krater typisch. Nur sehr große Krater (ab etwa 30km Durchmesser) resultieren aus Einschlägen mit ausreichender Energie, um Schmelz- und damit Fließ- oder Schwapp-Prozesse relevant für ihre Endform zu machen. Sie sind durch einen abgeflachten Boden und einen zentralen Gipfel gekennzeichnet. Auch ist generell das Alter nur solcher großer Krater hoch genug, um durch Weltraum-Erosion merklich gezeichnet zu sein. Durch diese Prozesse wird ein Krater von seinem tiefsten Punkt herauf mit Regolith-Auswurf naher weiterer Einschläge angefüllt oder seine Strukturen werden durch solche Einschläge teilweise durchbrochen oder überlagert. Wood und Anderson definierten

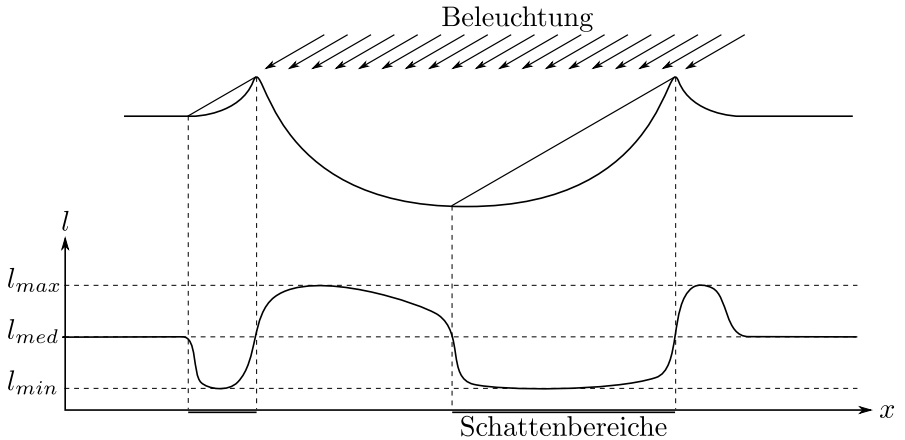


Abbildung 2.1: Schnitt durch das physische Profil eines typischen Mondkraters (oben) unterlegt mit dem aus der eingezeichneten Beleuchtung resultierenden Helligkeitsprofil (unten) *Bildquelle* (Maass, 2010)

eine nützliche Kategorisierung der anzutreffenden Kratertypen (Wood und Anderson, 1978), anhand von Beispielen in Abb. 2.2 illustriert.

Abb. 2.3 illustriert am Beispiel des Kraters CLAVIUS, wie sich die einzelnen Kategorien überlagern können, abhängig von ihrem Alter und ihrer Größe. Der Hauptkrater selbst misst ca. 225km im Durchmesser und ist mit seinen größten Nebenkraftern PORTER und RUTHERFURD (beide Durchmesser ca. 52km) vom Typ TYC. Halbkreisförmig von RUTHERFURD aus setzt sich eine charakteristische Kraterkette über die namenlosen Nebenkraftern D, C, N und J fort, die mit abnehmendem Durchmesser (28km bis 12km) in die Kategorien SOS bis BIO fallen.

Die Heuristik des Kratererkennungsalgorithmus ist mit den abgebildeten Typen ALC, BIO und SOS im Sinn formuliert, da diese in Bildern ein sehr homogenes Aussehen zeigen. Es gibt insbesondere keine kleineren Strukturen im Kraterinneren, die durch unerwartete Licht- und Schattengebilde den Algorithmus stören könnten. Das ist aber nur insofern eine Einschränkung, als dass das Verfahren für Krater oberhalb von 30km Durchmesser potentiell nicht geeignet ist. Solche Skalen sind für kein offensichtliches Navigationsszenario relevant, man müsste bei einer Bildbreite von 1024px und einem Öffnungswinkel von  $35^\circ$  eine Entfernung von 5000km vom Krater überschreiten, damit er unter einen abgebildeten Durchmesser von 10px schrumpft. Gängige Landeszenarien

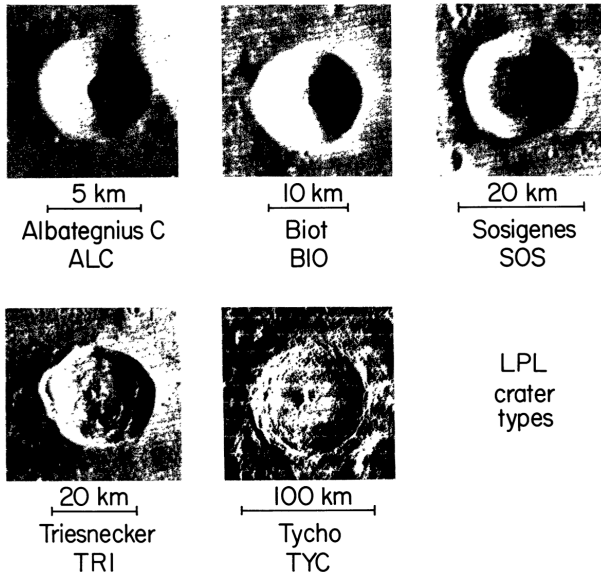


Abbildung 2.2: Fünf charakteristische morphologische Kategorien von lunaren Kratern  
*Bildquelle* (Wood und Anderson, 1978)

gehen von einem Navigationsbeginn in wenigen 10km bis 100km Höhe über dem lunaren Referenzradius aus.

Viele auch kleinere Krater sind dennoch unregelmäßig, etwa aufgrund von überschneidenden weiteren Einschlügen oder des Profils des ursprünglichen Untergrundes, auf dem der Krater entstand. Auf welche Weise solche Unregelmäßigkeiten zum Versagen von auf Kantendetektion basierenden Kratererkennungsverfahren führen, da sie als Modellannahme unter anderen die Konvexität der Kraterränder voraussetzen, wurde in (Maass, Krüger und Theil, 2011) erörtert. Das hier verwendete Erkennungsverfahren ist so ausgelegt, dass es unempfindlicher gegen genau diese Unregelmäßigkeiten ist als die kantendetektierenden und -verarbeitenden Verfahren.

Abb. 2.4 zeigt anhand eines Bildausschnittes eines typischen aber unregelmäßigen Kraters die charakteristischen Hell-Dunkel-Zonen in seinem Inneren, in Verbindung mit der lokal vorherrschenden Beleuchtungsrichtung. Man erkennt auch, dass alle in dem Bildausschnitt sichtbaren Krater bezüglich dieser Beleuchtungsrichtung die Anordnung ihrer dunklen und hellen Bereiche

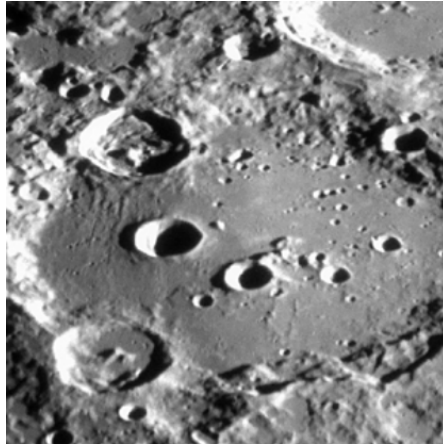


Abbildung 2.3: Morgenaufnahme des Kraters CLAVIUS auf der erd zugewandten Seite des Mondes mit Nebenkratern RUTHERFURD (oben links) und PORTER (unten links) *Bildquelle* Conrad Jung, Chabot Space & Science Center

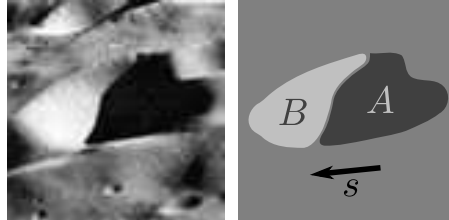


Abbildung 2.4: Charakteristische Hell-Dunkel-Zonen in einem unregelmäßigen Krater der Kategorie ALC mit dem Vektor der lokal dominanten Beleuchtungsrichtung *Bildquelle* (Maass, Krüger und Theil, 2011)

teilen: Die dunklen liegen stets vor den hellen Gebieten mit Bezug auf die Beleuchtungsrichtung.

Diese lokale charakteristische Anordnung der Kontrastgebiete in enger Nachbarschaft, verbunden mit der globalen robusten Ausrichtung an der realen Beleuchtungsrichtung, liefert ein mächtiges Kriterium zur Bewertung der „Kraterartigkeit“ eines betrachteten Bildausschnittes.

### 2.1.2 Implementierung

Die Innovation des Kratererkennungsverfahrens ist die Methode der Extraktion dieser Hell-und-Dunkel-Bereiche aus dem Bild. Die Schwierigkeit liegt hierbei darin, dass in typischem unvorprozessiertem Bildmaterial die Begriffe „hell“ bzw. „dunkel“ nur relative Eigenschaften beschreiben, sie haben nur in Bezug auf den mittleren Helligkeitswert ihrer direkten Umgebung Bedeutung. Die nächstliegende Lösung zu ihrer Isolierung, eine klassische Segmentierung durch Schwellwertbildung, beinhaltet damit ein großes Risiko: dass ein signifikanter Anteil aller ROIs unsichtbar bleibt, dadurch, dass viele von ihnen ganz oder teilweise unterhalb oder oberhalb einer falsch gewählten Schwelle liegen. Die in diesem Sinne optimale Schwelle war in Abb. 2.1 gekennzeichnet als  $l_{\text{med}}$ .

Das Bild des in Abb. 2.4 links gezeigten Kraters zerfällt unter der oben besprochenen Schwellwertbildung zum hypothetischen idealen Helligkeitswert genau in die rechts skizzierten Gebiete *A* und *B*. Die Komplexität des Problems verbirgt sich also hinter der Wahl des zur Zerlegung verwendeten Schwellwerts. Bild-globale Helligkeitsdurchschnitte wären beispielsweise ungeeignet aufgrund der oben besprochenen Variabilität der Helligkeitsverteilung über ein Bild, das möglicherweise sonnenzu- und sonnenabgewandte Teile einer großen Geländestruktur zeigt oder auch stark texturierten Grund enthält. Rein lokale Betrachtungen sind andererseits schwer, da hier das „Henne-und-Ei“-Problem auftritt, dass man die lokale Umgebung ja gerade erst finden möchte, über die man die Schwelle ermitteln müsste.

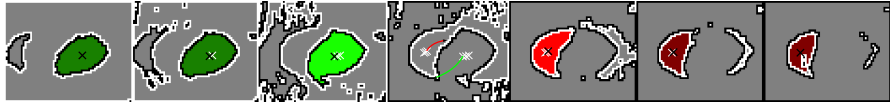
Die Lösung dieses Problems fand sich in einer Beobachtung über das Verhalten von zusammenhängenden Gebieten in einem Schwellwertbild zu einem dynamischen Schwellwert, dass nämlich beim Anheben dieses Schwellwertes die zusammenhängenden Flächen des Kraterinneren stetig anwachsen und dann bei Erreichen der mittleren Helligkeit der lokalen Umgebung eine Unstetigkeit aufweisen. Die Stetigkeit dieses Wachstumsprozesses ließ sich über die Bewegung des Schwerpunktes der Gebiete als stellvertretende Eigenschaft algorithmisch fassen.

Abb. 2.5a zeigt einen Bildausschnitt mit einem Krater, zu dem über sieben Helligkeitsschritte das jeweilige Schwellwertbild erzeugt wurde. Zur Isolation des Schattenbereiches der rechten Hälfte des Kraters wird im linken Feld der Abbildung begonnen und im korrespondierenden Feld der Abb. 2.5b das entsprechende zusammenhängende Gebiet markiert (dunkelgrün). Bei Anhebung des Schwellwertes im rechts-nächsten Feld wächst dieser Bereich an, bleibt aber isoliert und sein Schwerpunkt bewegt sich dabei nur marginal. Genauso geschieht es beim Übergang vom zweiten auf das dritte Feld von





(a) Anhebung des Schwellwertes und jeweils resultierendes Binärbild



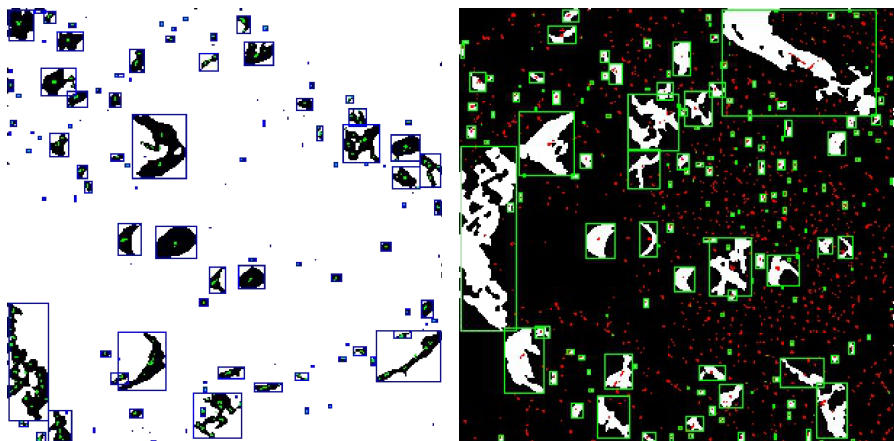
(b) Isolierte Gebiete in den einzelnen Schwellwertbildern zur obigen Zerlegung

Abbildung 2.5: Laufende Schwellwertzerlegung eines Bildausschnittes von CLAVIUS-D mit stabilen Gebieten nach dem Kriterium stetiger Schwerpunktbewegung  
*Bildquelle* (Maass, Krüger und Theil, 2011)

links: Der neue Schwerpunkt verbleibt in einer kleinen Umgebung des bzw. der Vorherigen. Im vierten Schritt dann erreicht im mittleren Feld der Schwellwert die lokale Umgebungshelligkeit und der isolierte Bereich bricht schlagartig auf. Hier verschiebt sich der Schwerpunkt dann unstetig, angedeutet durch den grünen Pfeil.

Die Isolierung des hellen Bereiches im linken Teil des Kraters verläuft analog vom rechten Feld aus linkswärts, wobei der Schwellwert sukzessive abgesenkt wird, bis im mittleren Feld auch dieses Gebiet nicht länger isoliert ist und der Schwerpunkt springt (roter Pfeil). Der hellgrüne Bereich im dritten Feld von links und der hellrote Bereich im dritten Feld von rechts decken in ihrer Vereinigung jetzt das Kraterinnere ab und bilden damit theoretisch genau die ROI, nach der gesucht wurde. Die Hauptkomponentenellipse über alle Pixel dieser Vereinigung der Gebiete ist dann ein robuster Deskriptor, der den Rand des Kraters im Bild approximativ parametrisiert.

Was dabei aber noch gar nicht vorliegt, ist das Wissen darüber, dass genau dieses hellrote Gebiet zu genau diesem hellgrünen Gebiet „gehört“. Betrachtet man in Abb. 2.6 das Ergebnis des vorgestellten Verfahrens zum gesamten Bild, aus dem der oben behandelte Ausschnitt stammte, ist diese Zuordnung auch nicht sofort offensichtlich. Dem Verfahren sind bloß die beiden Mengen der in Abb. 2.6a und 2.6b abgegrenzten Gebiete mit ihren Schwerpunkten bekannt. Kontext über irgendwelche physischen Ursprünge der ermittelten Werte ist nicht verfügbar. Zur Zeit der Entwicklung des Verfahrens wurde dies darüber aufgelöst, dass die externe Beleuchtungsrichtung als bekannt vorausgesetzt war und damit in dieser Beleuchtungsrichtung die beiden Gebiete von allen



(a) Stabile dunkle Gebiete (blau eingefasst) mit Schwerpunkten (grün)      (b) Stabile helle Gebiete (grün eingefasst) mit Schwerpunkten (rot)

Abbildung 2.6: Zerlegung der CLAVIUS-Bildes (Abb. 2.3) in seine stabilen Hell- und-Dunkel-Gebiete  
*Bildquelle* (Maass, 2010)

möglichen Zuordnungen den geringsten Abstand haben mussten.

Dieses Wissen über die Beleuchtungsrichtung vorauszusetzen, wird unmöglich gemacht durch die Erweiterung des reinen Krater-Detektionsproblems, das Gegenstand der Diplomarbeit des Autors war, auf die Zielstellung, ohne jegliches Vorwissen über die Pose der Kamera zu navigieren. Daher wurde im Rahmen dieser Dissertation eine Methode entwickelt, aus der Gesamtverteilung der stabilen Gebiete im obigen Sinne eine global dominante Beleuchtungsrichtung zu schätzen. Diese Methode ist Gegenstand des Abschnitts 2.2.

## 2.2 Erweiterungen im Rahmen der Dissertation

Das Resultat der in Abb. 2.6 gezeigten Zerlegung ist eine Liste von Deskriptoren zusammenhängender Gebiete des Trägers des Bildes, die im oben beschriebenen Sinne stabil waren über einen großen Bereich von Schwellwerten. Im folgenden Unterabschnitt werden diese Begriffe formal definiert.

## 2.2.1 Begriffsdefinitionen

Zu allererst muss erklärt werden, was ein „Bild“ im Sinne der hier entwickelten Algorithmen ist.

**Definition 2.1 (Bild)** Es seien  $m$ ,  $n$  und  $h$  natürliche Zahlen. Die diskreten Mengen  $\Omega$  und  $H$  mit

$$\begin{aligned}\Omega &:= \{1, \dots, m\} \times \{1, \dots, n\} \subset \mathbb{N}^2 \text{ und} \\ H &:= \{0, \dots, h-1\} \subset \mathbb{N}_0\end{aligned}$$

werden Träger bzw. Wertebereich genannt. Die zwischen diesen Mengen operierende Abbildung

$$I : \Omega \rightarrow H$$

heißt Bild. Sie ist nur an den Punkten des Trägers  $\Omega$  definiert und offensichtlich beschränkt. Zu jedem Bild ist punktweise  $I'$  definiert mit

$$I'(p) := h - 1 - I(p) \text{ für alle } p \in \Omega$$

und wird das inverse Bild zu  $I$  genannt.

Dieser Bildbegriff orientiert sich an der Notation aus (Bredies und Lorenz, 2010). Zur Abb. 2.5 war in Worten der Algorithmus der Extraktion intensitäts-stabiler zusammenhängender Gebiete beschrieben worden. Dabei wurde der Stabilitätsbegriff über die kleine Relativbewegung der Schwerpunkte dieser Gebiete bei ansteigendem Schwellwert erklärt. Auch diese Begriffe von Zusammenhang und Stabilität lassen sich formal fassen.

**Definition 2.2 (Stabile Gebiete)** Auf dem Träger  $\Omega$  eines Bildes  $I$  mit Wertebereich  $H$  gemäß Def. 2.1 seien die folgenden Eigenschaften von Gebieten definiert:

a) **Umgebung** Zu jedem Punkt  $p \in \Omega$  bezeichnet die Menge

$$U_{\Omega}(p) := \left\{ p' \in \Omega \mid \sum_{i \in \{1,2\}} |p'_i - p_i| \leq 1 \right\}$$

die 1-Umgebung von  $p$  in  $\Omega$ .

b) **Zusammenhang** Jede Untermenge  $\Omega' \subseteq \Omega$  heißt zusammenhängendes Gebiet, wenn es für je zwei Punkte  $p, q \in \Omega'$  mindestens eine Folge von

### 2.2.1 Begriffsdefinitionen

Punkten  $(p_i)_{1 \leq i \leq k}$  der Länge  $k$  gibt mit  $p_1 = p$ ,  $p_k = q$  und

$$\forall i : p_{i+1} \in U_{\Omega}(p_i)$$

Die Menge aller in diesem Sinne zusammenhängender Gebiete sei bezeichnet durch

$$\mathcal{T}_{\Omega} := \{\Omega' \subseteq \Omega \mid \Omega' \text{ ist zusammenhängend}\}.$$

**c) Stabilität** Eine Folge von Gebieten  $(G_k^-)_{1 \leq k \leq s}$  in  $\mathcal{T}_{\Omega}$  heißt stabil dunkel vom Grad  $s$  bezüglich des Bildes  $I$ , wenn für alle  $k = 1, \dots, s$  gilt

- 1)  $G_k^- \subset G_{k+1}^-$
- 2)  $G_{k+1}^- \setminus G_k^- \neq \emptyset$
- 3)  $\min\{I(\delta p) \mid \delta p \in G_{k+1}^- \setminus G_k^-\} > \max\{I(p) \mid p \in G_k^-\}$
- 4)  $\left\| 1/|G_{k+1}^-| \sum_{p \in G_{k+1}^-} p - 1/|G_k^-| \sum_{p \in G_k^-} p \right\|_2 \leq tol_{\text{stab}}$

Der Ausdruck  $|G_k^-|$  bezeichnet dabei die Anzahl der Punkte des Trägers in  $G_k^-$ . Eine Folge von Gebieten  $G_k^+$  heißt stabil hell vom Grad  $s$  und wird mit dem Exponenten „+“ versehen, wenn sie stabil dunkel vom Grad  $s$  bezüglich des invertierten Bildes  $I'$  ist.  $tol_{\text{stab}}$  ist als Abstandsschwelle der Schwerpunkte ein freier Parameter.

Jede Folge von stabil dunklen bzw. hellen Gebieten wird durch ihr jeweils letztes (gebietsweise größtes) Glied identifiziert. Da es ein großer Aufwand von Rechnerressourcen wäre, zu jedem dieser Gebiete die Menge aller zur Gebietsfolge gehörenden Punkte explizit zu speichern, wird darauf aufbauend noch ein Deskriptor definiert.

**Definition 2.3 (Deskriptor stabiler Gebiete)** Jede vom Grad  $s$  stabile dunkle oder stabil helle Folge  $(G_k^{\pm})_{1 \leq k \leq s}$  im Sinne der Def. 2.2c („±“ kann für „+“ oder „-“ stehen) sei eindeutig identifiziert über den aus Kenntnis von  $s$  und aus ihrem letzten Glied  $G_s^{\pm}$  abgeleiteten Deskriptor

$$S(G_s^{\pm}) := (s, n^{\pm}, c^{\pm}, P^{\pm})$$

mit

$$n^{\pm} := |G_s^{\pm}|, \quad c^{\pm} := \sum_{p \in G_s^{\pm}} p \quad \text{und} \quad P^{\pm} := \sum_{p \in G_s^{\pm}} pp^{\top}.$$

Die Menge dieser Deskriptoren aller mindestens vom Grad  $s$  stabil dunklen bzw.

stabil hellen Gebiete sei zusammengefasst im jeweiligen Deskriptorenvektor

$$\mathbf{S}^\pm := (S_1^\pm, S_2^\pm, \dots, S_{N^\pm}^\pm)$$

mit

$$S_i^\pm = S(G_{s_i}^\pm) = (s_i, n_i^\pm, c_i^\pm, P_i^\pm)$$

für alle Folgen  $(G_{k_i}^\pm)_{1 \leq k_i \leq s_i}$ ,  $i = 1, \dots, N^\pm$ , die stabil dunkel bzw. stabil hell vom Grad  $s_i \geq s$  sind.  $N^-$  und  $N^+$  bezeichnet die Anzahl stabil dunkler bzw. stabil heller Gebiete im Bild.

Diese Deskriptoren enthalten gerade alle notwendigen Elemente, die zur Aufstellung des Zuordnungsgraphen und später zur Bildung des Deskriptors der Krater-ROIs notwendig sind. Außerdem haben sie den Vorteil, dass sie beim Iterieren der Gebietsfolgen inkrementell aufgebaut werden können. D.h. das Hinzufügen von Punkten zu einem Gebiet und das Vereinigen von Gebieten ist trivial in dem Sinne, dass die Deskriptorelemente der Größe, des Schwerpunktes, und der Kovarianzmatrix schlicht addiert werden können.

Mit diesen Konstruktionen kann man jetzt zur Abb. 2.6 auf Seite 28 zurückkehren. Die in der Unterabbildung 2.6a eingerahmten Gebiete kennzeichnen die jeweils letzten Glieder stabil dunkler Gebietsfolgen, die in Unterabbildung 2.6b gekennzeichneten die letzten Glieder der stabil hellen Folgen. Auch wenn diese Gebiete noch punktweise gekennzeichnet sind in den Abbildungen, stehen den Algorithmen nur die aus ihnen abgeleiteten Deskriptorvektoren  $\mathbf{S}^-$  und  $\mathbf{S}^+$  zur Verfügung. Darauf aufbauend soll jetzt das Verfahren vorgestellt werden, das im Rahmen dieser Arbeit zur Generierung der Krater-ROIs über die Schätzung der Beleuchtungsrichtung entwickelt wurde.

## 2.2.2 Verfahren zur Beleuchtungsschätzung

Die Deskriptorvektoren  $\mathbf{S}^-$  und  $\mathbf{S}^+$  enthalten jeweils Elemente, die Kontrastgebiete beschreiben, die entweder nicht Teil von abgebildeten Kratern sind oder die keine Entsprechung im jeweils anderen Deskriptorenvektor besitzen. Es handelt sich deswegen nicht um eine eindeutige Zuordnung von Elementen des einen zu Elementen des anderen und sie sind generell auch von unterschiedlicher Länge. Daher muss aus allen möglichen Zuordnungen eine in irgendeinem Sinne optimale gefunden werden.

Wie am Ende des Abschnitts 2.1.2 besprochen wurde, kann von der grundlegenden Annahme ausgegangen werden, dass die Untermenge der korrekten Zuordnungen der Deskriptorvektoren zueinander in Bezug auf die dominante

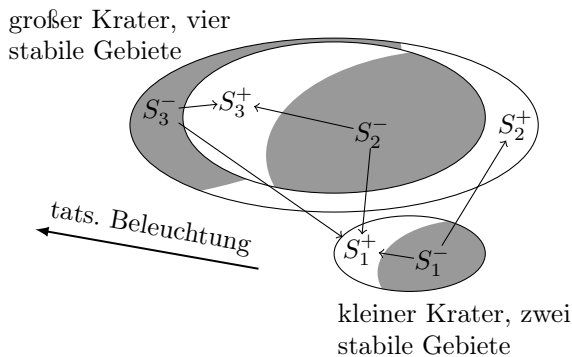


Abbildung 2.7: Größenabhängige unterschiedliche Anzahlen von dunklen und hellen Gebieten pro Krater und vollständiger Zuordnungsgraph nächster und nächstnächster Partnergebiete

Beleuchtungsrichtung die Abstände der Schwerpunkte der einander zugeordneten Deskriptoren minimiert. Je nach der Beleuchtungselevation über dem Gelände und der abgebildeten Größe eines Kraters können aber seine äußeren Hänge ebenfalls beschienen bzw. schattiert sein und daher stabil helle bzw. dunkle Gebiete auf dem Bildträger erzeugen.

Das führt dazu, dass ein stabil dunkles Gebiet im Inneren des Kraters einem stabil hellen Gebiet am äußeren Hang des Kraters zugeordnet werden könnte, oder umgekehrt, und diese Zuordnung sich auch nicht ohne Kenntnis der Beleuchtungsrichtung ausschließen lässt. Unter Umständen kann zusätzlich der Abstand der Schwerpunkte zweier zusammengehörender stabiler Gebiete von dem eines dritten, sehr kleinen, Gebietes in einem benachbarten Krater unterschritten werden.

Abb. 2.7 illustriert dies und verdeutlicht, wie man aus diesen Gründen zu jedem stabil dunklen Gebiet das jeweils nächste und nächstnächste stabil helle Gebiet zur Zuordnung in Betracht ziehen muss. Diese potentielle Doppelzuordnung erzeugt einen einfachen, ungewichteten, gerichteten Graphen:

**Definition 2.4 (Nachbarschaftsgraph stabiler Gebiete)** Gegeben seien die Deskriptorenlisten  $S^-$  und  $S^+$  nach Def. 2.3. Dazu sei die  $N^- \times N^+$ -Ab-

standsmatrix  $(d_{ij}^{\pm})$  definiert mit

$$d_{ij}^{\pm} := \left\| \frac{c_i^-}{n_i^-} - \frac{c_j^+}{n_j^+} \right\|_2$$

für je ein  $S_i^-$  aus  $\mathcal{S}^-$  und  $S_j^+$  aus  $\mathcal{S}^+$ . Zu jedem  $i$ -ten stabil dunklen Gebiet, für  $i = 1, \dots, N^-$ , definieren die jeweils kleinsten und zweitkleinsten Elemente seiner Zeile in dieser Abstandsmatrix die  $i$ -te Zeile  $A_i = (a_{i1}, a_{i2})$  der  $N^- \times 2$ -Adjazenzliste  $A$  des Graphen über

$$\begin{aligned} a_{i1} &:= \arg \min_j \{d_{ij}^{\pm} \mid 1 \leq j \leq N^+\} \\ a_{i2} &:= \arg \min_j \{d_{ij}^{\pm} \mid j = 1, \dots, a_{i1} - 1, a_{i1} + 1, \dots, N^+\} \end{aligned}$$

Der Nachbarschaftsgraph ist vollständig definiert über diese Adjazenzliste: Der Zeilenindex ist der Index des Knotens im Deskriptorenvektor  $\mathcal{S}^-$  und seine Zeilenelemente in  $A$  zeigen an, zu welchen Knoten in  $\mathcal{S}^+$  Kanten von ihm ausgehen. Der Graph in Abb. 2.7 wäre nach dieser Definition beispielsweise definiert über

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ 3 & 1 \end{pmatrix} \quad (2.1)$$

Da alle Kanten des Graphen orientiert sind von einem dunklen Knoten zu einem hellen Knoten, wird jetzt die Annahme relevant, die Kanten „korrekter“ Zuordnungen würden im Mittel mit der dominanten Beleuchtungsrichtung zusammenfallen. Die dominante Beleuchtungsrichtung im Bild kann am einfachsten als Winkel über seine Zeilen-Achse ausgedrückt werden. Jede der gerichteten Kanten des Graphen lässt sich auf diesen Winkel abbilden durch Berechnung der Verbindungslinie der im Deskriptor verzeichneten Schwerpunkte.

Für einen gegebenen Graphen  $A$  nach Def. 2.4 wird die Menge  $\Theta$  dieser Winkel erzeugt durch Ziehen aller  $2N^-$  Kanten zwischen den Schwerpunkten stabil dunkler Gebiete zu den durch die Adjazenzliste angezeigten Schwerpunkten stabil heller Gebiete. Die Abbildung

$$\theta(i, l, A) := \text{atan2} \left( \left( \frac{c_{a_{il}}^+}{n_{a_{il}}^+} - \frac{c_i^-}{n_i^-} \right)^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \quad (2.2)$$

für  $c_{a_{il}}^+, n_{a_{il}}^+ \in S_{a_{il}}^+$  und  $c_i^-, n_i^- \in S_i^-$ , liefert zu einem Index  $i$  des Deskrip-

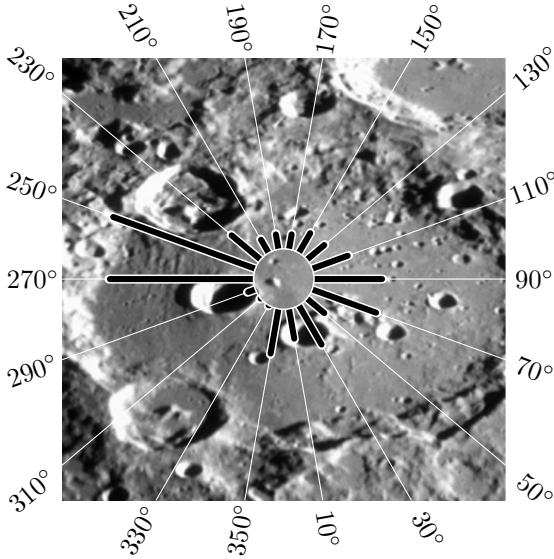


Abbildung 2.8: Krater CLAVIUS überlagert mit dem Richtungshistogramm der Kanten des Nachbarschaftsgraphen aller stabil dunklen und hellen Gebiete

torenvektors der stabil dunklen Gebiete und einem Index  $l$  den Winkel der aus der Adjazenzliste des Graphen und durch das Indexpaar  $i$ - $l$  eindeutig bestimmten Kante. Die Menge der Winkel aller Kanten des Graphen wird gebildet durch

$$\Theta := \{\theta(i, l, A) \mid i = 1, \dots, N^-, l = 1, 2\}. \quad (2.3)$$

Die Information über die dominante Beleuchtungsrichtung lässt sich aus einem Histogramm der Winkelmenge  $\Theta$  ableiten. Abb. 2.8 zeigt das Bild des Kraters CLAVIUS aus Abb. 2.3 überlagert mit einem Histogramm von  $\Theta$  ( $20^\circ$  pro Intervall).

Der aus einem Bild wie dem Beispiel abgeleitete Nachbarschaftsgraph enthält oft nur eine relativ geringe Menge von Kanten und damit von Winkeln. Einfach ein Histogramm-Maximum über  $\Theta$  zu suchen, wäre daher empfindlich gegenüber zufälligen Spitzen in der Verteilung der Winkel, je nach Quantisierung des Histogramms. Deswegen wird zur Schätzung der Beleuchtungsrichtung eine Gütefunktion definiert, die diesen Problemen gegenüber robust ist.



**Definition 2.5 (Schätzfunktion der Beleuchtungsrichtung)** Zur Winkelmenge  $\Theta$  nach (2.3) zu einem Graphen  $A$  der Nachbarschaften stabil dunkler und heller Gebiete eines Bildes  $I$  sei die Einschränkung auf ein Fenster der Breite  $\pi$  um einen mittleren Wert  $\bar{\theta}$  definiert als

$$\Theta|_{\bar{\theta}} := \left\{ \theta \in \Theta \mid -\frac{\pi}{2} \leq (\theta \circ \bar{\theta}) \leq \frac{\pi}{2} \right\}$$

mit der „Differenz modulo  $2\pi$ “

$$\theta \circ \bar{\theta} := \begin{cases} \theta - \bar{\theta} & : |\theta - \bar{\theta}| \leq \pi \\ -2\pi + \theta - \bar{\theta} & : \theta - \bar{\theta} > \pi \\ 2\pi + \theta - \bar{\theta} & : \theta - \bar{\theta} < -\pi \end{cases}$$

Die Abbildung  $\phi$  der gesamten Winkelmenge

$$\phi(\Theta) := \arg \min_{\bar{\theta} \in [0, 2\pi)} \frac{1}{|\Theta|_{\bar{\theta}}} \sum_{\theta \in \Theta|_{\bar{\theta}}} (\theta \circ \bar{\theta})^2$$

auf den Punkt, über den das zweite statistische Moment im um ihn gelegten  $\pm \frac{\pi}{2}$ -Fenster minimal ist, wird die Schätzfunktion der Beleuchtungsrichtung genannt.  $|\Theta|_{\bar{\theta}}$  steht dabei für die Anzahl der Winkel in  $\Theta|_{\bar{\theta}}$ .

$\phi$  bestimmt über jeden Punkt in  $[0, 2\pi)$  das zweite statistische Moment in einem gleitenden Halbkreis-Fenster auf der Verteilung in  $\Theta$  und liefert genau den Winkel zurück, an dem dieses Minimum auftritt. Illustriert ist dies im Diagramm 2.9, wo zusätzlich zum Histogramm aus dem Beispielbild der Verlauf der Schätzfunktion und das Minimum  $\bar{\theta}$  gezeigt wird.

Der Grund für die Konstruktion dieser speziellen Funktion liegt wieder in der an Abb. 2.7 gemachten Beobachtung, dass große Krater u.U. bis zu vier stabile Gebiete erzeugen können, und dass es drei verschiedene Möglichkeiten ihrer Zuordnung zueinander gibt. Zwei dieser Zuordnungen sind dann irrtümlich, es wird je ein stabiles Gebiet am äußeren Hang des Kraters mit einem Gebiet in seinem Inneren verbunden. Die dadurch implizierte Beleuchtungsrichtung ist dann annähernd um  $180^\circ$  gedreht zur tatsächlichen. Tatsächlich prägen sich in der Praxis Winkelverteilungen  $\Theta$  aus, die sich wie die Summe zweier gewickelter Normalverteilungen verhalten, mit annähernd  $180^\circ$  versetzten Mittelwerten und unterschiedlichen Standardabweichungen.

Da man aus den o.g. Gründen als Abstand der Mittelwerte genau diesen halben Kreis erwarten kann, lässt sich aber die Komplexität der Bestimmung

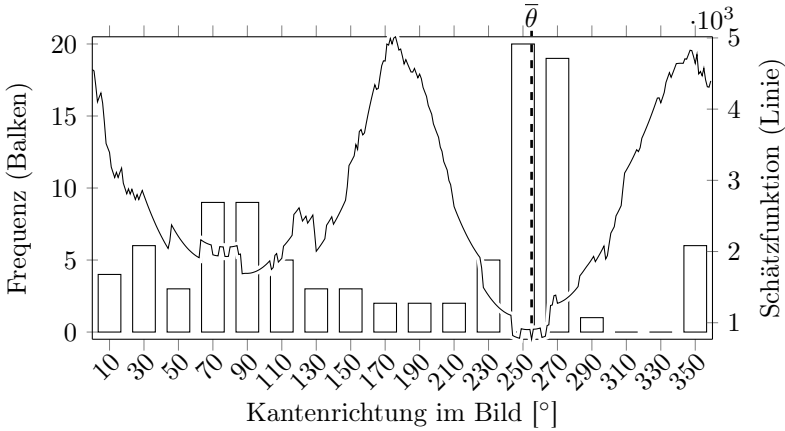


Abbildung 2.9: Richtungshistogramm der Kanten des Nachbarschaftsgraphen aus Abb. 2.8, Verlauf von  $\sigma$  aus (2.4) und Minimum der Schätzfunktion der Beleuchtungsrichtung aus Def. 2.5

der Parameter einer solchen Summe von Verteilungen vernachlässigen. Indem man das gleitende Fenster von genau einem Halbkreis wählt, schließt man in seiner Betrachtung am Optimum den Großteil des Beitrags der „gegenüberliegenden“ Verteilung zur Gesamtwahrscheinlichkeitsmasse im Fenster aus. Die Verteilung der im Fenster verbleibenden Winkel ist dann am Optimum ungefähr normal, und das Ergebnis der Schätzfunktion an diesem Punkt approximiert den Punkt minimaler Varianz der Winkel.

Diese Approximation der Varianz an dem durch  $\phi$  bestimmten optimalen Winkel  $\bar{\theta}$  sei definiert über

$$\sigma^2 := \frac{1}{|\Theta|_{\bar{\theta}}} \sum_{\theta \in \Theta|_{\bar{\theta}}} (\theta \circ \bar{\theta})^2. \quad (2.4)$$

Die Menge  $\Theta|_{\bar{\theta}}$  und der Operator  $\circ$  waren in Def. 2.5 eingeführt worden. Das oben beschriebene Verhalten lässt sich am Beispielhistogramm in Abb. 2.8 bzw. Abb. 2.9 ablesen: Es existieren die zwei erwarteten Spitzen, nahezu um  $180^\circ$  versetzt.

$\sigma$  lässt sich im Sinne einer „Standardabweichung“ als Kriterium zum Ausschluss von irrtümlichen Graphenkanten verwenden, indem alle Kanten entfernt werden, deren Winkel um mehr als ein fest gewähltes Vielfaches von

$\sigma$  vom mittleren Winkel  $\bar{\theta}$  abweichen. Über diesen Ausschluss wird schließlich die Menge der ROIs definiert:

**Definition 2.6 (Deskriptoren der Krater-ROIs)** Zum Bild  $I$  sei der Nachbarschaftsgraph stabil heller und dunkler Gebiete  $A$  bekannt. Der minimale Wert der Approximation ihrer Varianz (2.4) an der Stelle  $\bar{\theta}$  wird durch  $\sigma$  bezeichnet.

**a) Reduzierter Graph** Zu  $A$  bezeichne  $A' = (a'_{il})$  den auf  $\bar{\theta} \pm \sigma$  reduzierten Graphen mit

$$a'_{il} := \begin{cases} a_i : & |\theta(i, l, A) \ominus \bar{\theta}| \leq \sigma \\ 0 : & \text{sonst} \end{cases}, \quad i = 1, \dots, N^-, \quad l = 1, 2,$$

mit der Winkelfunktion  $\theta(i, l, A)$  aus (2.2).

**b) Vereinigung der Gebiet-Deskriptoren** Ist ein Element  $a'_{il}$  des reduzierten Nachbarschaftsgraph  $A'$  nicht Null, so bildet es auf den  $j$ -ten gemeinsamen Schwerpunkt  $c_j$  und die gemeinsame Kovarianz  $P_j$  der referenzierten Deskriptoren ab durch

$$c_j := \frac{c_i^- + c_{a'_{il}}^+}{n_i^- + n_{a'_{il}}^+} \quad \text{und} \quad P_j := \frac{P_i^- + P_{a'_{il}}^+}{n_i^- + n_{a'_{il}}^+},$$

wobei  $j$  die Anzahl der zeilenweise gezählten bisherigen nicht-Null-Elemente von  $(a'_{il})$  ist.

**c) Deskriptor der Krater-ROI** Zu jeder Kovarianzmatrix  $P_j$  bezeichne  $\lambda_{j,1}$  den größeren bzw.  $\lambda_{j,2}$  den kleineren Eigenwert, und  $z_{j,1}$  den Eigenvektor zum größeren Eigenwert bzw.  $z_{j,11}$  sein erstes Element. Dann ist das Tupel  $E_j = (x_j, y_j, a_j, b_j, \theta_j)$ , definiert durch

$$E_j := \begin{pmatrix} x_j = c_j \\ y_j \\ a_j = \sqrt{2\lambda_{j,1}} \\ b_j = \sqrt{2\lambda_{j,2}} \\ \theta_j = \text{atan2} \left( \text{sign}(z_{j,11}) z_{j,1} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \end{pmatrix}^T,$$

der Deskriptor der  $j$ -ten Region of Interest eines Kraters.  $E_j$  parametrisiert eine Ellipse und approximiert die Abbildung des physischen Randes des

Kraters. Im Sonderfall eines Kreises, wenn  $P_j$  zwei gleichgroße Eigenwerte besitzt, ist der Winkel  $\theta_j$  im Deskriptor frei wählbar.

**d) Deskriptorenvektor aller ROIs** Der Vektor von Tupeln  $E_j$

$$\mathbf{E} := (E_1, \dots, E_{N_p})$$

mit  $N_p \leq 2N^-$  definiert die Deskriptorliste aller solcher ROIs.

Die Matrix  $P_j$  aus Def. 2.6b ist als akkumulierte Kovarianz der Gebietspunkte symmetrisch und positiv definit. Daher sind die in Def. 2.6c verwendeten Eigenwerte  $\lambda_{j,1}$  und  $\lambda_{j,2}$  reell und größer Null. Der Winkel  $\theta_j$  bewegt sich zwischen  $-\pi$  und  $\pi$ , er repräsentiert anschaulich die kleinste (positive oder negative) Rotation der Ellipse, mittels derer sich ihre große Halbachse auf die Hochachse des Bildes drehen lässt.

## 2.3 Zwischenfazit

Der bildverarbeitende Teil des in dieser Arbeit vorgestellten Gesamtverfahrens ist damit bereits abgeschlossen. An der Schnittstelle zur nächsten Verarbeitungsstufe wird der Deskriptorenvektor  $\mathbf{E}$  nach Def. 2.6d übergeben. In ihm sind diejenigen Ellipsen enthalten, von denen der Algorithmus vermutet, dass sie die physischen Ränder abgebildeter Krater approximieren. Das Eingangsbild  $I$  wird in den nachfolgenden Kapiteln nicht mehr benutzt.

Abhängig von der Abbildungsqualität kann es nützlich sein, die Ellipsen  $\mathbf{E}$  vor der Weiterverarbeitung noch einer „Verfeinerung“ zu unterziehen, in deren Rahmen die Ellipse morphologisch auf ein lokales Gradientenmaximum im Bild eingepasst wird. Eine rechnerisch günstige und effektive Methode hierzu wurde in (Maass, Krüger und Theil, 2011) vorgestellt. Es soll aber an dieser Stelle nicht weiter darauf eingegangen werden, denn es zeigte sich über die Jahre in einer Vielzahl von Test- und Anwendungsfällen, dass dieser Schritt nur in einem sehr engen Bereich äußerer Parameter (Beleuchtungseigenschaften, Geländetypen, Kamerapose) eine mehr als marginale Verbesserung bietet und daher im Sinne einer Reduktion der Komplexität des Verfahrens auch weggelassen werden kann.

Während die im Abschnitt 2.1 beschriebene grundlegende Methode der Kraterdetektion bereits im Jahr 2011 funktionierte, ist die im Abschnitt 2.2 gemachte Erweiterung für das Funktionieren des Gesamtverfahrens dieser Arbeit ein unverzichtbarer Schritt, denn sie macht die Detektionsmethode unabhängig von der Kenntnis äußerer Bedingungen. Das war eine der Grund-

voraussetzungen dafür, dass ein Raumfahrzeug mit einem auf diesem Verfahren basierenden Sensor sich völlig ohne vorheriges Wissen über seine Lage oder Position orientieren kann.

Mit den hier geschaffenen Voraussetzungen wird jetzt im folgenden Kapitel der Übergang zur Modellbildung gemacht. Aus kontextfreien Informationen in Form der Deskriptorenliste  $\mathbf{E}$  müssen jetzt Modelle konstruiert werden, die nicht nur innerhalb der Menge der detektierten Krater Verhältnisse wie relative Größen und Entfernungen abzuleiten vermögen, sondern letztendlich eine Korrespondenzbildung zu einem globalen Katalog von bekannten Kratern erlauben.



# Kapitel 3

## Startwertlose Krateridentifikation

### 3.1 Einleitung und Motivation

Nach Abschluss der bildverarbeitenden Stufe des Gesamtverfahrens bildet dieses Kapitel den theoretischen Kernteil der vorliegenden Arbeit. Es ist daher auch das umfangreichste. Die im Nachsatz zum letzten Kapitel genannte Modellbildung kann durch eine kurz formulierte Problemstellung motiviert werden, die das auf Seite 7 formulierte Gesamtziel dieser Dissertation präzisiert.

„Zu einem aus der Anwendung des Kratererkennungsverfahrens auf ein Eingabebild resultierenden Satz von Ellipsen und einem vorhandenen Katalog von bekannten global verorteten Kratern, identifiziere ohne weiteres Vorwissen eine ausreichende Menge der beobachteten Krater, um die Kamerapose ableiten zu können, aus der das Bild aufgenommen wurde.“

Mit Bezug auf diese Zielstellung folgt ein kurzer Ausblick auf die Abschnitte dieses Kapitels und ihre Zusammenhänge untereinander:

Zunächst werden in Abschnitt 3.2 die verwendeten Kamera- und Bildkoordinatensysteme definiert, die den Übergang vom zweidimensionalen Bild auf den dreidimensionalen Raum ermöglichen, in dem sich die aufnehmende Ka-

mera bewegt. Daraus wird in Abschnitt 3.3 zu den Bildellipsen der Satz ihrer projektiven elliptischen Kegel rekonstruiert und es wird untersucht, wie und unter welchen Annahmen aus ihnen dann wiederum auf die dreidimensionalen Parameter der abgebildeten Kraterländer geschlossen werden kann.

Die Abschnitte 3.4 bis 3.6 enthalten Grundlagendefinitionen der modellbasierten Beschreibung der beobachteten Mondoberfläche und erläutern Rahmenbedingungen und Voraussetzungen zum gewählten Modellierungsansatz. Mit ihrer Hilfe und aus den in Abschnitt 3.3 rekonstruierten Parametern der dreidimensionalen Kraterbeschreibung wird im Abschnitt 3.7 ein konkreter Algorithmus hergeleitet, um ein Freiform-Modell der in einem Bild sichtbaren Mondoberfläche zu rekonstruieren. Verbleibende Mehrdeutigkeiten bezüglich der Lösungen der dabei auftretenden Gleichungen werden durch Konstruktion von geeigneten Heuristiken im Abschnitt 3.8 aufgelöst.

Der Abschnitt 3.9 liefert schließlich die Methoden der Korrespondenzbildung zwischen in einem Bild detektierten Kratern und dem mitgeführten globalen Kraterkatalog. Auch das Format dieses Katalogs wird dort eingeführt.

Es werden während der folgenden Konstruktionen immer auch Hintergrundüberlegungen dokumentiert, um den „roten Faden“ von den vorausgesetzten Daten zu den abstrakten Modellen zu ziehen, die die Lösung des zentralen Problems dieser Dissertation bestimmen. Teilweise ist es notwendig, nicht offensichtliche Annahmen kurz zu belegen, da auf ihnen die Lösbarkeit der konstruierten Systeme fußt. Dies wird aber auf ein Minimum begrenzt. Zu einigen Konstruktionen, die sehr abstrakt bleiben, sind einige einfache Rechenbeispiele zur Erläuterung im Anhang A verfügbar. Im Zweifel sind Definitionen eher ausführlicher gehalten, um ein komplettes und konsistentes Rahmenwerk bereitzustellen, das aus sich heraus die Implementierung des Verfahrens erlaubt. Konstruktionen, die sich besser sequentiell als deklarativ beschreiben lassen, sind durch Pseudocode-Algorithmen erläutert.

## 3.2 Lochkameramodell und Kamerakoordinatensysteme

Die Aufnahme von Bildern durch physische Kameras lässt sich immer approximativ modellieren, um einen ungefähren projektiven Zusammenhang zwischen beobachteten dreidimensionalen Objekten vor der Kamera herzustellen mit ihren Abbildungen auf eine Abbildungsfläche (z.B. ein Film oder Bildsensor). Oft wird in der Literatur zur robotischen Bildverarbeitung ein sehr allgemeines Abbildungsmodell eingeführt, das eine Vielzahl von Parame-



tern offen lässt, wie Brennweite, Bildhauptpunkt, Pixel-Rechteckigkeit und mehrere Verzerrungsparameter eines nichtlinearen Bildverzerrungsmodells zur Modellierung von Unzulänglichkeiten der Kamera und ihres Linsenapparates. Dieses komplexe Thema wird in dieser Arbeit nicht angeschnitten werden. Es wird zur Vereinfachung stets angenommen, dass diese Parameter im Vorhinein geschätzt und herauskorrigiert wurden. Diese Vereinfachung ist insofern zulässig, als dass alle für die Verfahrenstests in Kapitel 5 verwendeten künstlich erzeugten Bilder ohnehin dieser Bedingung genügen und für alle mit realen Kameras aufgenommenen Bilder durch sorgfältige Kalibration und vorherige Entzerrung dieser Zustand hergestellt wird.

Was verbleibt, ist der einfache projektive Zusammenhang der Beziehung zwischen Bild- und Kamerakoordinatensystemen, wie dargestellt in Abb. 3.1. Das Bild selbst liegt in den „Matrixkoordinaten“ der Definition 2.1 aus

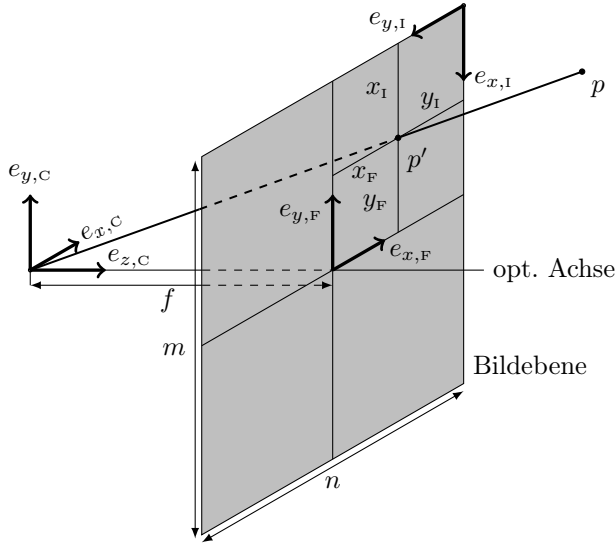


Abbildung 3.1: Kamera- und Bildkoordinatensysteme

Abschnitt 2.2 vor, d.h. in Tupeln  $u_I = (x_I, y_I)^T \in \Omega$  je zweier natürlicher Zahlen (Subskript I, für *Image*). Es wird dabei bezüglich des dargestellten Bildes  $x_I$  positiv von oben nach unten und  $y_I$  positiv von links nach rechts gezählt. Die Koordinatenachsen  $e_{x,I}$  und  $e_{y,I}$  sind also in diesen Richtungen in der linken oberen Ecke des Bildes festgemacht.

### 3.3.0 Räumliche Rekonstruktion der Kraterränder

---

Um von Bildkoordinaten  $(e_{x,I}, e_{y,I})$  zu Kamerakoordinaten  $(e_{x,C}, e_{y,C}, e_{z,C})$  zu konvertieren, wird das „Zwischensystem“ des Bildrahmens  $(e_{x,F}, e_{y,F})$  genutzt (Subskript F, für *Frame*). Mit der Bildhöhe  $m$  und Bildbreite  $n$  in Pixeln ist der Zusammenhang zwischen Bildkoordinaten und Bildrahmenkoordinaten

$$\begin{pmatrix} x_F \\ y_F \end{pmatrix} = \frac{s}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} y_I \\ x_I \end{pmatrix}, \quad s = \max\{m, n\}. \quad (3.1)$$

Die Beziehung (3.1) ist natürlich umkehrbar, also eine Transformation. Dies gilt jedoch nicht mehr für den Übergang von Bildrahmenkoordinaten zu Kamerakoordinaten oder dessen Umkehrung. Dies ist ersichtlich, da es sich bei der Abbildung eines räumlichen Punktes  $p$  auf seinen Bildpunkt  $p'$  auf der Bildebene um eine perspektivische Projektion handelt (vgl. Abb. 3.1). Die Beziehung dort ist:

$$\begin{pmatrix} x_F \\ y_F \end{pmatrix} = \frac{f}{z_C} \begin{pmatrix} x_C \\ y_C \end{pmatrix}. \quad (3.2)$$

Die skalare Größe  $f$  ist hierbei die Brennweite der Kamera. Man sieht jedoch bereits, dass (3.2) nicht eindeutig umkehrbar ist. Von einem Punkt  $p'$  auf der Bildebene ist nur die Gerade bekannt, auf der er im Raum vor der Kamera liegt, die dritte Koordinate bleibt frei.

## 3.3 Räumliche Rekonstruktion der Kraterränder

Kapitel 2 stellt das Verfahren zur Verfügung, das aus einem Bild der beleuchteten Mondoberfläche Krater erkennt und ihre Ränder als Bildellipsen  $\mathbf{E}$  approximiert als Ergebnis zurückliefert. Aus diesen Ellipsen müssen die dreidimensionalen Parameter ihrer jeweilig zugrundeliegenden Krater abgeleitet werden. Dazu sei ein Krater modelliert als ein Kreis  $C$  im Raum, vollständig definiert durch seinen Mittelpunkt  $c$ , seine Oberflächennormale  $n$  und seinen Radius. Die perspektivische Projektion dieses Kreises durch das oben eingeführte Lochkameramodell auf eine Bildebene ergibt unter diesem Modell dann wieder die Ellipse  $E$ .

Alle vorliegenden Ellipsen  $E \in \mathbf{E}$  in Bildrahmenkoordinaten, mit Parametern

$$E = (x_F, y_F, a, b, \theta), \quad (3.3)$$

sind zunächst möglicherweise die Projektion des räumlichen Kreises  $C$  eines Kraterrandes auf die Bildebene. Die Rekonstruktion dieses Kreises aus der Ellipse ist nicht völlig trivial, man muss wie folgt vorgehen:

1. Bestimmung des elliptischen Kegels  $K$  dessen Schnitt mit der Bildebene die Ellipse beschreibt und durch den Brennpunkt läuft
2. Schnitt des Kegels mit einer Ebene  $P$ , so, dass die Schnittkurve ein Kreis  $C$  ist
3. Bestimmung des Mittelpunktes  $c$  des Kreises und der Ebenennormale  $n$

Die genannten geometrischen Figuren sind in der Abbildung 3.2 mit ihren projektiven Zusammenhängen illustriert.

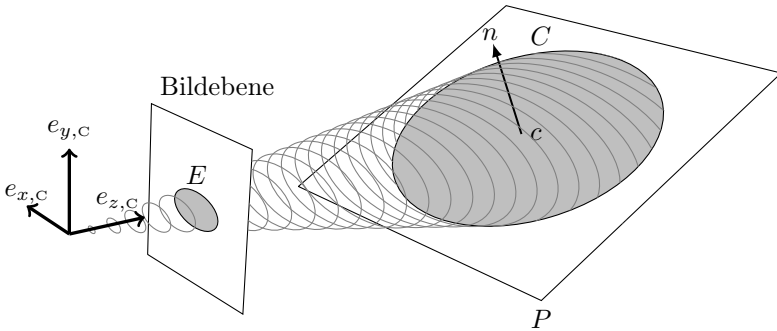


Abbildung 3.2: Zusammenhang eines räumlichen Kreises mit einer Bildellipse über den projektiven Kegel

Die oben genannten Berechnungen, ausgeführt in den folgenden Abschnitten, konnten vom Autor in einer mit der vorliegenden Anwendung kompatiblen Art nicht gefunden werden. Sie wurden daher von den obigen Annahmen ausgehend selbst hergeleitet und sind der Vollständigkeit halber entsprechend ausführlich beschrieben.

### 3.3.1 Projektive Kegel der Kraterellipsen

Gegeben sei die Ellipse  $E$  in Bildrahmenkoordinaten, gesucht die Matrix der quadratischen Funktion  $K$  desjenigen Kegels, den die Projektion aller ihrer Randpunkte erzeugt. Dann gilt für jeden Punkt  $u = u_C \in \mathbb{R}^3$  im Kamerakoordinatensystem

$$u^T K u = 0, \quad (3.4)$$

### 3.3.1 Projektive Kegel der Kraterellipsen

---

wobei  $K$  selbst die Form

$$K = \begin{pmatrix} k_1 & k_4 & k_5 \\ k_4 & k_2 & k_6 \\ k_5 & k_6 & k_3 \end{pmatrix} \quad (3.5)$$

hat. Gleichung (3.4) ist ausgeschrieben für  $u = (x, y, z)^\top$

$$k_1 x^2 + k_2 y^2 + k_3 z^2 + 2(k_4 xy + k_5 xz + k_6 yz) = 0. \quad (3.6)$$

Die Gleichungen (3.4) bzw. (3.6) beschreiben dabei eine Quadrik, die bei passender Wahl der Parameter  $k_1$  bis  $k_6$  den gesuchten räumlichen Kegel durch den Ursprung des Kamerakoordinatensystems ergeben. Aus der Bildellipse  $E$  kann dieser Kegel relativ leicht gewonnen werden, denn die Bestimmung der sechs Parameter in  $K$  kann wegen (3.6) als die Lösung des Gleichungssystems

$$\begin{pmatrix} x_1^2 & y_1^2 & z_1^2 & 2x_1y_1 & 2x_1z_1 & 2y_1z_1 \\ x_2^2 & y_2^2 & z_2^2 & 2x_2y_2 & 2x_2z_2 & 2y_2z_2 \\ x_3^2 & y_3^2 & z_3^2 & 2x_3y_3 & 2x_3z_3 & 2y_3z_3 \\ x_4^2 & y_4^2 & z_4^2 & 2x_4y_4 & 2x_4z_4 & 2y_4z_4 \\ x_5^2 & y_5^2 & z_5^2 & 2x_5y_5 & 2x_5z_5 & 2y_5z_5 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \\ k_6 \end{pmatrix} = 0 \quad (3.7)$$

interpretiert werden für fünf Punkte  $u_1$  bis  $u_5$ , die auf der Ellipse  $E$  liegen. Die Koordinaten der  $u_i$  liegen hierbei wieder in Kamerakoordinaten vor und lassen sich über die Parametrisierung der Ellipse  $E$  gewinnen, etwa aus

$$u_{i,C} = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} x_F + a \cos \varphi_i \cos \theta - b \sin \varphi_i \sin \theta \\ y_F + a \cos \varphi_i \sin \theta + b \sin \varphi_i \cos \theta \\ 1 \end{pmatrix}, \quad i = 1, \dots, 5. \quad (3.8)$$

für  $\varphi_i = i2\pi/5$ . Weil die Ellipse  $E$  auf der Bildebene liegt und daher nach (3.8) für alle fünf  $u_i$  die  $e_{z,C}$ -Koordinate Eins ist, wird das System (3.7) daher vereinfacht zu

$$\underbrace{\begin{pmatrix} x_1^2 & y_1^2 & 1 & 2x_1y_1 & 2x_1 & 2y_1 \\ x_2^2 & y_2^2 & 1 & 2x_2y_2 & 2x_2 & 2y_2 \\ x_3^2 & y_3^2 & 1 & 2x_3y_3 & 2x_3 & 2y_3 \\ x_4^2 & y_4^2 & 1 & 2x_4y_4 & 2x_4 & 2y_4 \\ x_5^2 & y_5^2 & 1 & 2x_5y_5 & 2x_5 & 2y_5 \end{pmatrix}}_{=:M} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \\ k_6 \end{pmatrix} = 0. \quad (3.9)$$

Die Matrix  $M$  hat nach Konstruktion (3.8) nun Rang 5 und der gesuchte Parametervektor  $k := (k_1, \dots, k_6)^T$  spannt ihren Kern auf:

$$\text{rang}(M) = 5 \quad \text{und} \quad \ker(M) = k \quad (3.10)$$

Die Berechnung des Kerns von  $M$  ist über gängige numerische Methoden sehr effizient möglich, etwa lässt er sich über die redundante Spalte der  $Q$ -Matrix einer QR-Zerlegung von  $M^T$  numerisch stabil erzeugen.

### 3.3.2 Kamerafeste Oberflächennormalen der Krater

Die Matrix  $K$  der quadratischen Funktion des Kegels ist damit bekannt. Sei nun die Ebene  $P$  (vgl. Abb. 3.2), mit der der Kegel geschnitten werden soll, parametrisiert durch

$$p \in P : p = q + \xi r + \eta s, \quad q, r, s \in \mathbb{R}^3, \quad \xi, \eta \in \mathbb{R}, \quad (3.11)$$

wobei  $q$  der Stützvektor und  $r, s$  die Erzeugendenvektoren der Ebene sind (es muss natürlich gelten  $r \neq \lambda s$  für alle  $\lambda \in \mathbb{R}$ ). Da für beliebige Vektoren  $a$  und  $b$  und eine symmetrische Matrix  $A$  gilt,

$$(a + b)^T A (a + b) = a^T A a + 2a^T A b + b^T A b, \quad (3.12)$$

außerdem für einen Skalar  $\alpha$ ,

$$(\alpha a)^T A (\alpha a) = \alpha^2 a^T A a, \quad (3.13)$$

ergibt Einsetzen von  $p$  aus (3.11) in (3.4)

$$\begin{aligned} 0 &= p^T K p = (q + \xi r + \eta s)^T K (q + \xi r + \eta s) \\ &= q^T K q + (\xi r)^T K (\xi r) + (\eta s)^T K (\eta s) \\ &\quad + 2(q^T K (\xi r) + q^T K (\eta s) + (\xi r)^T K (\eta s)) \\ &= r^T K r \xi^2 + s^T K s \eta^2 + q^T K q \\ &\quad + 2(q^T K r \xi + q^T K s \eta + r^T K s \xi \eta). \end{aligned} \quad (3.14)$$

Dies ist aber wieder eine quadratische Funktion in homogenen Koordinaten  $\xi$  und  $\eta$  für eine Quadrik in zwei Dimensionen und lässt sich in Matrixform schreiben:

$$p^T K p = \begin{pmatrix} \xi & \eta & 1 \end{pmatrix} \underbrace{\begin{pmatrix} r^T K r & r^T K s & q^T K r \\ r^T K s & s^T K s & q^T K s \\ q^T K r & q^T K s & q^T K q \end{pmatrix}}_{=: K'} \underbrace{\begin{pmatrix} \xi \\ \eta \\ 1 \end{pmatrix}}_{=: p'} = p'^T K' p' = 0 \quad (3.15)$$


---

### 3.3.2 Kamerafeste Oberflächennormalen der Krater

Für Kegelschnittgleichungen der Art (3.15) kann man allgemeine Aussagen über ihre Klassifizierung machen: Sei  $D$  die Determinante von  $K'$ ,  $d$  die Determinante der linken oberen  $2 \times 2$ -Matrix von  $K'$  und  $m$  die Spur der linken oberen  $2 \times 2$ -Matrix von  $K'$ . Der Kegelschnitt (3.15) ist genau dann eine Ellipse, wenn  $d > 0$  und  $mD < 0$  ist. Damit er sogar ein Kreis ist, müssen noch

$$\begin{aligned} r^T K r - s^T K s &= 0 \quad \text{und} \\ r^T K s &= 0 \end{aligned} \tag{3.16}$$

erfüllt sein. Die Bedingungen (3.16) haben zwar keine offensichtlichen Lösungen  $r$  und  $s$ , die folgende Konstruktion führt aber zum gewünschten Ergebnis:

Seien  $z_1, z_2, z_3$  die Eigenvektoren der Kegelschnittmatrix  $K$ , jeweils zu den Eigenwerten  $\lambda_1, \lambda_2, \lambda_3$ . Für solche Kegelschnittmatrizen gilt entweder

$$\begin{aligned} \lambda_1 &\geq \lambda_2 > 0 > \lambda_3 \quad \text{oder} \\ \lambda_3 &> 0 > \lambda_2 \geq \lambda_1, \end{aligned} \tag{3.17}$$

je nach Vorzeichen von  $K$ . Dann steht  $z_1$  im Raum orthogonal auf der großen Halbachsenebene des Kegels (diejenige Ebene, die durch die großen Halbachsen der elliptischen Schnitte des Kegels mit den zu seinem Mittelstrahl orthogonalen Ebenen aufgespannt wird). Genauso steht  $z_2$  orthogonal auf die kleine Halbachsenebene und  $z_3$  spannt genau den Mittelstrahl des Kegels auf. Sei jetzt  $q := z_3$  und  $r := z_2$  fest gewählt, wie in Abbildung 3.3 illustriert.

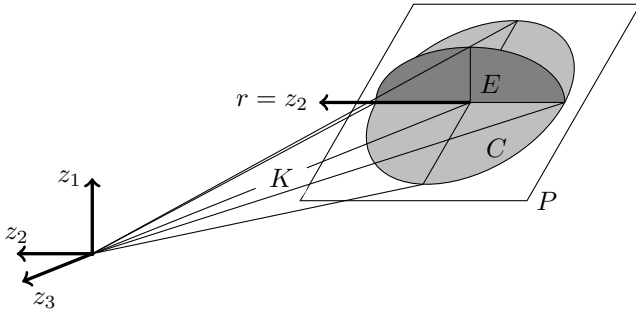


Abbildung 3.3: Herstellung eines kreisartigen Kegelschnittes durch Rotation einer Ebene über einen Eigenvektor der Kegelquadratik

Mit dieser Wahl von  $r$  lässt sich der zweite Vektor,  $s$ , bestimmen. Mit der Überlegung, dass der gesuchte Vektor  $s$   $K$ -orthogonal zu  $r$  stehen soll (die



### 3.3.3 Kreisförmige Kegelschnitte

---

$$\alpha^2 + \beta^2 = 1 \Leftrightarrow \beta^2 = 1 - \alpha^2:$$

$$\begin{aligned}\alpha^2 &= \frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3} \\ \Leftrightarrow \alpha_{1,2} &= \pm \sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} \\ \Rightarrow \beta_{1,2} &= \sqrt{1 \mp \frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}}\end{aligned}\tag{3.21}$$

Man sieht auch, warum  $\alpha$  und  $\beta = \sqrt{1 - \alpha^2}$  immer reell sein müssen: Wegen (3.17) folgt, dass entweder  $\lambda_1 - \lambda_3 \leq \lambda_2 - \lambda_3 < 0$  oder  $\lambda_1 - \lambda_3 \geq \lambda_2 - \lambda_3 > 0$  sind und damit auch ihr Quotient größer Null und kleiner oder gleich Eins ist. Wählt man jetzt einfach die positive Lösung von  $\alpha$ , sind damit die beiden Vektoren  $r$  und  $s$  bekannt, und der gesuchte Normalenvektor  $n$  der Ebene  $P$  ist durch ihr Kreuzprodukt gegeben:

$$r = z_2, \quad s = z_1 \sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} + z_3 \sqrt{1 - \frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}}\tag{3.22}$$

und

$$n = r \times s.\tag{3.23}$$

### 3.3.3 Kreisförmige Kegelschnitte

Der Kreis  $C$  ist die Quadrik, die durch die Matrix  $K'$  in (3.15) beschrieben wird. Da bereits klar ist, dass  $p'^T K' p' = 0$  einen Kreis beschreibt, gelten die Bedingungen (3.16). „Normiert“ man dann die Matrix  $K'$  durch Division mit  $s^T K s$ ,

$$\frac{1}{s^T K s} K' = \begin{pmatrix} 1 & 0 & \frac{q^T K r}{s^T K s} \\ 0 & 1 & \frac{q^T K s}{s^T K s} \\ \frac{q^T K r}{s^T K s} & \frac{q^T K s}{s^T K s} & \frac{q^T K q}{s^T K s} \end{pmatrix},\tag{3.24}$$

und vergleicht mit der grundsätzlichen Form einer Kegelschnittmatrix, die einen Kreis beschreibt,

$$\begin{pmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ -x_0 & -y_0 & x_0^2 + y_0^2 - r^2 \end{pmatrix},\tag{3.25}$$



(wobei  $x_0$  und  $y_0$  die Mittelpunktkoordinaten und  $r$  der Radius sind), kann man die gesuchte Größe direkt ablesen. Der Einheitsvektor  $c$ , der dann in Richtung des Kreismittelpunktes von  $C$  zeigt, ist also gegeben durch

$$c = \frac{q - s \frac{q^T K s}{s^T K s}}{\left\| q - s \frac{q^T K s}{s^T K s} \right\|}. \quad (3.26)$$

### 3.3.4 Mehrdeutigkeit der Lösungen

Der Normalenvektor ist nicht eindeutig bestimmt, denn es existieren zu jeder Ebene mit Normale  $n$  drei weitere gültige Lösungen, die einen Kreis im Schnitt mit  $P$  erzeugen. Sie sind durch die vier möglichen Vorzeichenkombinationen in (3.21) gegeben, lassen sich anschaulicher aber auch durch Vorzeichenwechsel und/oder Spiegelung an der Ebene finden, die auf den Mittelstrahl  $q$  des Kegels orthogonal steht. Die Lösungen sind dann  $n$ ,  $-n$ ,  $n' = n - 2qq^T n$  und  $-n'$ , vgl. Abb. 3.5. Zwei der Lösungen scheiden sofort aus,  $-n$  und  $n'$ ,

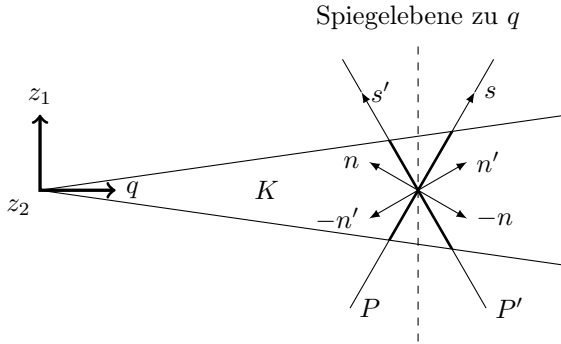


Abbildung 3.5: Mehrere Lösungen für den Normalenvektor der Schnittebene

denn zu  $n$  beschreibt  $-n$  dieselbe Ebene mit Erzeugendenvektor  $s$ , ebenso  $-n'$  zu  $n'$  und  $s'$ . Es lässt sich also ohne Einschränkungen das Vorzeichen so wählen, dass  $n$  und  $-n'$  (daher ab hier  $n'$ , ohne Minus) in Richtung der Kamera zeigen (denn die Alternative würde bedeuten, dass „von unten“ auf den Krater geschaut wird).

Es verbleibt die Ambivalenz bezüglich der Spiegelung an der Ebene mit Normalenvektor  $q$ . Der Kreismittelpunkt  $c'$  für den Schnitt mit der Ebene  $P'$  ist analog zu  $c$  über (3.26) mit der über  $q$  gespiegelten Richtung  $s'$  (vgl. Abb.

3.5) zu ermitteln, doch ohne weitere Information ist an dieser Stelle nicht entscheidbar, ob das Paar  $(n, c)$  oder  $(n', c')$  die im Bild sichtbare Oberfläche besser beschreibt. Wie diese Information aus dem Kontext aller erkannten und rekonstruierten Krater ermittelt wird, ist Gegenstand des Abschnitts 3.8.

## 3.4 Formale Grundlagen zur Oberflächenapproximation

Die aus der oben gemachten Rekonstruktion vorliegende Schar von Kreisnormalenvektoren  $n$  (und  $n'$ ) der gefundenen Kreise  $C$  stellt mit ihren Einheitsrichtungen  $c$  (und  $c'$ ) Messungen der tatsächlich beobachteten Mondoberfläche dar. Es verbleibt zu jedem Krater ein unbekannter Skalierungsfaktor  $\lambda$ , der aus der perspektivischen Projektion auf die Bildebene resultiert.  $\lambda$  korrespondiert zur unbekannten Entfernung der Kamera zum Krater. Weiterhin bleibt eine Zweideutigkeit der Rekonstruktion in den Zentren- und Normalenvektoren.

Abbildung 3.6 macht diese Zweideutigkeit anschaulich: Es werden von einer Kamera im Orbit des Mondes drei Kraterellipsen erkannt, doch die Rekonstruktion der Krater-Mittelpunktsrichtungen und Normalen erzeugt für jeden Krater zwei (zulässige) Lösungen. Im Bild scheinen alle  $n_i$  die „korrekten“ Lösungen zu sein, wie auch in Abb. 3.7. Im Allgemeinen könnte die Ausgangskonfiguration  $n_i$  vs.  $n'_i$  aber beliebig gemischt sein. In einfachen

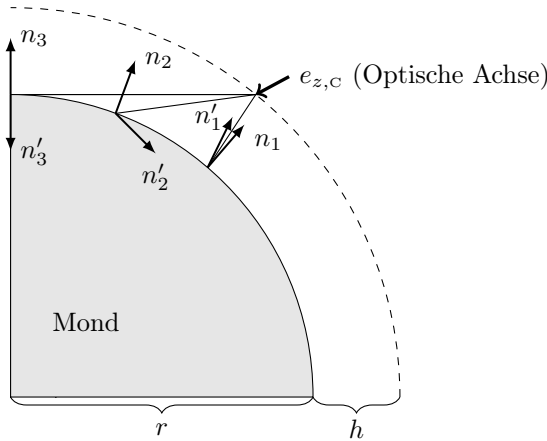


Abbildung 3.6: Doppelte Lösung  $n$  vs.  $n'$  für drei Krater.

Szenarien kann vermutet werden, welche der beiden Lösungen das Modell der Oberfläche unterstützt. Etwa, wenn bekannt ist, dass die Mondoerfläche im Bild „unten“ ist, also  $n$  nach „oben“ zeigen muss, oder umgekehrt. Im allgemeinen Fall aber, wenn die relative Position des Mondes nicht bekannt ist oder möglicherweise sogar die vollständige Krümmung der Oberfläche sichtbar ist (der Nadirvektor<sup>1</sup> zeigt ins Bild), hat man keine solchen hilfreichen Informationen.

Hier versteckt sich ein komplexes Problem: Bei  $k$  erkannten Kratern und ebensovielen Paaren  $(c, n)$ ,  $(c', n')$  ergeben sich schon  $2^k$  Kombinationsmöglichkeiten, die jeweils auf eine verschiedene Konfiguration der Mondoerfläche im Bild hinweisen. Die folgende Definition führt eine formale Notation ein, die es einfacher macht, über solche verschiedenen Konfigurationen der aus der Kratererkennung gewonnenen Oberflächendaten zu argumentieren:

**Definition 3.1 (Daten und Konfigurationen)** Sei  $N_p$  die Anzahl der erkannten Krater und  $0 < k \leq N_p$ .

**a) Daten** Eine Anzahl  $k$  wie in Abschnitt 3.3 rekonstruierter Zentren- und Normalenvektoren werden zusammengefasst im Tupel  $\mathbf{R}$  mit folgender Struktur:

$$\mathbf{R} := (R_1, R_2, \dots, R_k) = ((\mathbf{r}_1, \mathbf{r}'_1), (\mathbf{r}_2, \mathbf{r}'_2), \dots, (\mathbf{r}_k, \mathbf{r}'_k))$$

Hierbei bezeichnen  $R_{i,1} = \mathbf{r}_i := (c_i, n_i)$  bzw.  $R_{i,2} = \mathbf{r}'_i := (c'_i, n'_i)$  die zueinander symmetrischen Paare von räumlich rekonstruierten Zentrum- und Normalenvektoren zu einem erkannten Krater.

**b) Konfiguration** Der Indexvektor  $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_k) \in \{1, 2\}^k$  erzeugt eine Konfiguration der Daten durch die Projektion

$$R_i \mapsto R_{i,\kappa_i} := \begin{cases} (c_i, n_i) & \text{wenn } \kappa_i = 1 \\ (c'_i, n'_i) & \text{wenn } \kappa_i = 2 \end{cases}$$

Ein Datenpaar  $R_i$  wird auf dessen von  $\kappa_i$  angezeigte Komponente abgebildet. Das  $i$ -te durch  $\kappa_i$  konfigurierte Zentrum- und Normalenpaar wird gekennzeichnet mit dem Hut:  $(\hat{c}_i, \hat{n}_i) := R_{i,\kappa_i}$ . Die Kurzschreibweise  $\mathbf{R}^{[\boldsymbol{\kappa}]}$  beschreibt die

<sup>1</sup>Als Nadir wird hier der normierte Richtungsvektor vom Kamerakordinaten-Ursprung zum Mondmittelpunkt bezeichnet.

auf diese Weise konfigurierten Daten:

$$\mathbf{R}^{[\kappa]} = (R_{1,\kappa_1}, \dots, R_{k,\kappa_k}) = ((\hat{c}_1, \hat{n}_1), \dots, (\hat{c}_k, \hat{n}_k))$$

Es existieren  $2^k$  verschiedene  $k$ -Indexvektoren  $\kappa$  und daher genauso viele Konfigurationen der Daten  $\mathbf{R}$ .

**c) Symmetrische Konfiguration** Zu jeder Konfiguration  $\kappa$  bezeichnet  $\kappa'$  die symmetrische Konfiguration, wenn beide die selbe Länge  $k$  haben, und gilt

$$\kappa'_i = \begin{cases} 1 & \text{wenn } \kappa_i = 2 \\ 2 & \text{wenn } \kappa_i = 1 \end{cases}$$

für alle  $i = 1, \dots, k$ .

Die Konfiguration der Daten liefert nur die Stützwerte für die Oberflächenrekonstruktion. Zusammen mit einem Satz von Skalaren  $\lambda$ , der die relative Entfernung des Kraterzentrumspunktes in der jeweiligen Zentrumsrichtung kontrolliert, fehlt zur vollständigen Definition einer Oberfläche noch der folgende Glattheitsbegriff:

**Definition 3.2 (Rekonstruierte Oberfläche)** Sei  $\mathbf{R}$  ein Datensatz nach Def. 3.1a mit  $k$  Elementen und einer Konfiguration  $\kappa$  gemäß Def. 3.1b. Dazu sei  $\lambda := (\lambda_1, \dots, \lambda_k) \in (0, \infty)^k$  ein Vektor von Skalaren. Eine Fläche  $S_\kappa^\lambda \subset \mathbb{R}^3$  heißt zu  $\mathbf{R}^{[\kappa]}$  rekonstruierte Oberfläche, wenn für sie gilt:

- i)  $\lambda_i \hat{c}_i \in S_\kappa^\lambda$
- ii)  $S_\kappa^\lambda$  ist in  $\lambda_i \hat{c}_i$  mindestens einmal stetig differenzierbar.
- iii)  $\hat{n}_i$  ist an diesem Punkt Oberflächennormale von  $S_\kappa^\lambda$ .

Eine schematische 2D-Illustration möglicher resultierender Oberflächen mit Begriffen aus Def. 3.2 ist in Abb. 3.7 für drei Krater gegeben, für 3 der möglichen  $2^3 = 8$  Konfigurationen der Normalen  $n_i, n'_i$ .

Das Rekonstruktionsproblem lässt sich an Abb. 3.7 schön illustrieren: Ohne weitere Informationen lässt sich aus den Daten  $\mathbf{R}$  keine Entscheidung fällen, welche der rekonstruierten Oberflächen die „echte“ beschreibt. Die linke Oberfläche  $S_{\kappa_1}^{\lambda_1}$  ist konvex, könnte also z.B. eine vollständige Mondkugel repräsentieren, im Falle der Oberfläche  $S_{\kappa_2}^{\lambda_2}$  könnte der Fall vorliegen, in dem die Kamera in eine Senke oder in einen Krater hineinschaut. In der Oberfläche  $S_{\kappa_3}^{\lambda_3}$  ließe sich eine Hügelkette erkennen. Außerdem ist zu jeder Konfiguration  $\kappa_i$  genau eine Oberfläche  $S$  dargestellt, festgelegt durch Wahl der Skalare

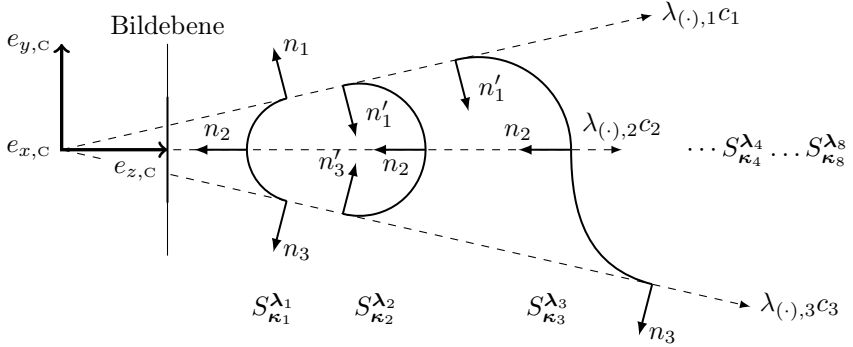


Abbildung 3.7: Verschieden konfigurierte und skalierte Oberflächen für drei Krater in 2D.

$\lambda_i$ .  $\lambda$  ist aber Parameter und es existiert daher eine ganze Klasse möglicher Oberflächen zu jeder Konfiguration  $\kappa$ , die der Definition 3.2 genügt.

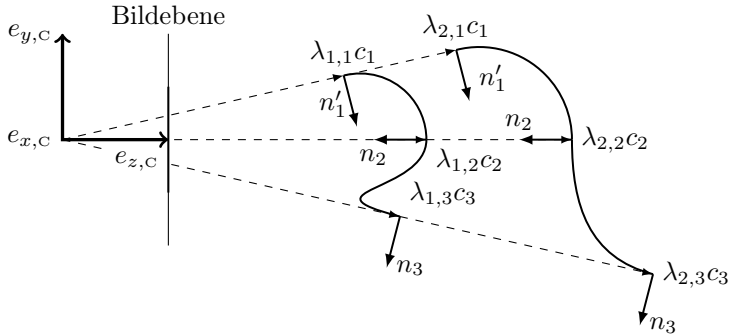


Abbildung 3.8: Zwei ähnliche Oberflächen

Dieses Ähnlichkeitsprinzip soll in Abb. 3.8 verdeutlicht werden. Dort sind zwei Oberflächen  $S^{\lambda_1}_{\kappa}$ ,  $S^{\lambda_2}_{\kappa}$  gezeigt mit selber Konfiguration  $\kappa$  der Daten aber mit unterschiedlichen relativen Skalierungen  $\lambda_1$  und  $\lambda_2 \approx (1.4, 1.5, 1.75) \cdot \lambda_1$ . Man erkennt die identische Konfiguration an der gleichen Orientierung der Normalen  $\hat{n}_i$  aber erhält große Skalen- und Krümmungsunterschiede.

Man sieht, die Anforderungen an die Oberflächen aus Definition 3.2 lässt die Details der tatsächlich verwendeten Modelle frei. In den folgenden Unter-

### 3.5.0 Parametrische Oberflächenmodelle

---

abschnitten werden zwei verschiedene Kategorien von Modellen eingeführt und teilweise an Beispielen erläutert, sowie Vor- und Nachteile diskutiert werden. Die zwei Arten von Modellen sind:

- I. Parametrische Modelle:** Durch eine kleine Anzahl Parameter vollständig bestimmte Formen im Raum. Vertreter sind alle Nullstellengebilde zweiter Ordnung wie Ellipsoide, Koniken, Ebenen, etc.
- II. Freie Modelle:** Nur teilparametrisch, etwa stückweise durch Polynome beschreibbar und genügen gewissen Übergangsbedingungen an den Stückgrenzen. Der in dieser Arbeit untersuchte Vertreter dieser Art von Modell wird eine Splinefläche sein.

Gruppe **I.** zeichnet sich durch Einfachheit aus. Es existieren simple Verfahren, die Modellparameter aus Linearen Ausgleichsproblemen über die Zentren- und Normalendaten der erkannten Krater zu erhalten. Dafür hängt die Nutzbarkeit der Lösung direkt von der Realitätstreue des Modells ab.

Gruppe **II.** ist wesentlich komplexer, sowohl in Darstellung und Auswertbarkeit, als auch dem Aufwand, ihre Parameter aus den Zentren- und Normalendaten zu gewinnen. Dafür haben diese Modelle das Potenzial, freie Formen mit vielfachen Krümmungswechseln treuer abzubilden.

## 3.5 Parametrische Oberflächenmodelle

Die Gruppe der parametrischen Modelle soll hier kurz angerissen werden, um die Unterschiede und Schwächen gegenüber dem später in dieser Arbeit verwendeten freien Modell herauszuarbeiten. Als Beispiel wird das vermeintlich intuitivste Modell für die Mondoberfläche gewählt: Die Kugel. Als geometrisches Gebilde ist sie vollständig durch ihren Mittelpunkt und ihren Radius bestimmt, es wird daher das Kugelmodell exemplarisch für die Kategorie **I.** besprochen.

Mit der Kugeleigenschaft, dass an jedem Punkt der Mondoberfläche die äußere Oberflächennormale radial vom Mondmittelpunkt nach außen zeigt, und dem bekannten mittleren Radius des Mondes ( $\approx 1.737.100\text{m}$ ), verbleibt als unbekannter zu schätzender Parameter der Mondmittelpunkt. Dieser muss aus den rekonstruierten Zentren- und Normalenvektoren  $c$  und  $n$  abgeleitet werden können.

Die rekonstruierten Normalen  $n$  sind normiert, d.h.  $\|n\|_2^2 = 1$ , für alle Krater, weswegen eine der drei Koordinaten redundant ist. Ihre Transformation

vom Kamerakoordinatensystem in ein Kugelkoordinatensystem  $(e_{\nu,C}, e_{\varphi,C})$  mit impliziter dritter Koordinate  $r_C \equiv 1$  ist daher bijektiv:

$$\begin{aligned}\nu_C &= \text{atan2}(y_C, x_C) \\ \varphi_C &= \arccos(z_C)\end{aligned}\tag{3.27}$$

Die  $\nu$ - $\varphi$ -Verteilung am Beispiel des Bildes vom Krater CLAVIUS (vgl. Abb. 2.3 auf Seite 25) ist in Abbildung 3.9 dargestellt. Die Breite-Länge-Darstellung

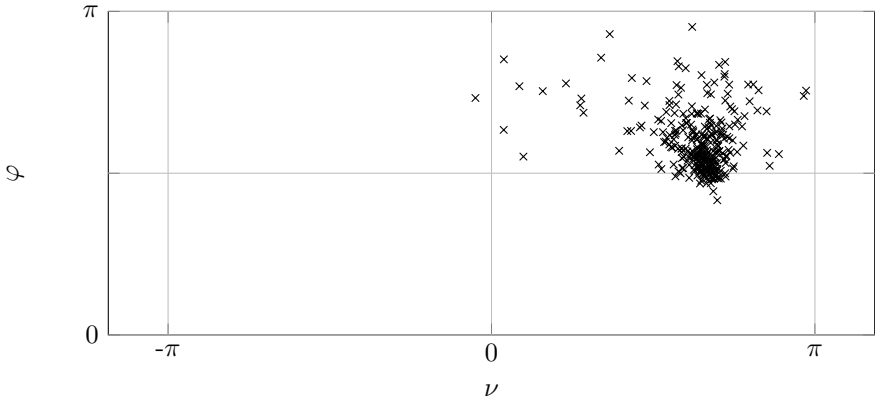


Abbildung 3.9: Verteilung der zwei ersten Kugelkoordinaten der rekonstruierten Normalen für ein Beispielbild

ist ganz offensichtlich nicht normalverteilt, daher genügt es nicht, das ungewichtete Mittel der Längen- und Breitenkoordinaten zu nehmen und dadurch auf den Nadir zurückzuschließen.

In der Literatur findet sich ein lineares Modell, um aus genau diesen Informationen die gesuchte Nadirrichtung zu ermitteln: Sei die Schar gefundener Krater definiert durch die Paare

$$c_i = (c_{i,1}, c_{i,2}, c_{i,3})^T \quad \text{und} \quad n_i = (n_{i,1}, n_{i,2}, n_{i,3})^T \tag{3.28}$$

ihrer Mittelpunkte bzw. Normalenvektoren im Kamerakoordinatensystem<sup>2</sup>. D. S. Wokes und S. M. Wokes stellten folgende Idee vor: Jeder Kratermittelpunkt  $c_i$  spannt mit dem Richtungsvektor  $m$  des Mondmittelpunkts und

<sup>2</sup>Die Zweideutigkeit  $(c, n)$  vs.  $(c', n')$ , die in den letzten Unterabschnitten behandelt wurde, ist durch die Struktur des Modells hier unerheblich.

### 3.5.0 Parametrische Oberflächenmodelle

dem Kameraursprung eine Ebene  $P_i$  auf, deren Normale durch  $n_{p,i} = c_i \times n_i$  gegeben ist. Im Kugelmodell ist die implizite Annahme enthalten, dass die Kraternormale radial vom Mondmittelpunkt nach außen zeigt. Die Gerade, die im Schnitt aller Ebenen  $P_i$  liegt, muss folglich die Nadir-Richtung  $m$  anzeigen.

Setzt man für die Ebene  $P_i$  die homogene Normalengleichung (denn der Ursprung liegt in der Ebene)

$$x^T n_{p,i} = 0, \quad x \in \mathbb{R}^3 \quad (3.29)$$

an, so muss gelten

$$m^T n_{p,i} = 0 \quad \text{für alle} \quad n_{p,i} = c_i \times n_i \quad (3.30)$$

und es ergibt sich das folgende Ausgleichsproblem für den Richtungsvektor vom Kameraursprung zum Mondmittelpunkt  $m := (m_1, m_2, m_3)^T$ :

$$\begin{aligned} &\text{minimiere} \quad \|Nm\|_2^2 \\ &\text{unter} \quad \|m\|_2^2 = 1. \end{aligned} \quad (3.31)$$

Die Matrix  $N$  aus (3.31) ist dabei die Matrix mit den Ebenennormalen  $n_{p,i}$  als  $i$ -te Zeile:

$$N := \begin{pmatrix} c_{1,2}n_{1,3} - c_{1,3}n_{1,2} & c_{1,3}n_{1,1} - c_{1,1}n_{1,3} & c_{1,1}n_{1,2} - c_{1,2}n_{1,1} \\ c_{2,2}n_{2,3} - c_{2,3}n_{2,2} & c_{2,3}n_{2,1} - c_{2,1}n_{2,3} & c_{2,1}n_{2,2} - c_{2,2}n_{2,1} \\ c_{3,2}n_{3,3} - c_{3,3}n_{3,2} & c_{3,3}n_{3,1} - c_{3,1}n_{3,3} & c_{3,1}n_{3,2} - c_{3,2}n_{3,1} \\ \vdots & \vdots & \vdots \end{pmatrix} \quad (3.32)$$

Das Problem (3.31) ist ein Minimierungsproblem mit einfacher Matrixnebenbedingung. Mithilfe des LAGRANGE-Multiplikators  $\lambda$  lässt sich die LAGRANGE-Funktion

$$L(m; \lambda) := \|Nm\|_2^2 - \lambda(\|m\|_2^2 - 1), \quad \lambda \neq 0 \quad (3.33)$$

definieren und die Lösung des Problems (3.31) wird zur Suche eines Minimums von (3.33). Dies wiederum ist äquivalent zur Suche nach den Nullstellen der Ableitung

$$\begin{aligned} \frac{d}{dm} L(m; \lambda) &= \frac{d}{dm} (\|Nm\|_2^2 - \lambda(\|m\|_2^2 - 1)) \\ &= \frac{d}{dm} (m^T (N^T N) m - \lambda(m^T m - 1)) \\ &= 2N^T N m - 2\lambda m \\ &\stackrel{!}{=} 0 \end{aligned} \quad (3.34)$$

$$\iff N^T N m = \lambda m,$$



was ein Eigenwertproblem der Matrix  $N^T N \in \mathbb{R}^{3 \times 3}$  beschreibt. Lösung ist der Nadirvektor  $m$ , mit dem sich dann die Flughöhe über dem mittleren Mondradius errechnen lässt (D. S. Wokes und S. M. Wokes, 2010). Das prinzipielle Vorgehen ist am Beispiel A.1, Anhang A demonstriert.

Die Lösung von (3.31) ist invariant bezüglich der beiden korrekten Lösungen  $n$  und  $n'$  der Kraterschar (vgl. Abb. 3.5). Dies ist insofern ein Vorteil, als noch nicht klar ist, wie man die Zweideutigkeit auflösen kann. D. S. Wokes und S. M. Wokes testeten ihr oben vorgestelltes Verfahren anhand synthetisch generierter Daten (Kugelmodell plus weißes Rauschen) und erhielten für moderates Rauschen i.A. vertretbare Fehler im Bereich weniger Grad für die Nadirrichtung. Treten wie im tatsächlichen Anwendungsfall jedoch systematische Fehler auf, etwa mit realen Resultaten einer Kratererkennung, wird der Null-Bruchpunkt des Kleinste-Quadrate-Ansatzes relevant und die ermittelte Lösung wird beliebig schlecht. Auch die Kugelannahme ist nur auf allergrößten (mondglobalen) Skalen vertretbar.

Der ultimative Nachteil dieses Ansatzes ist aber, dass ohne Auflösung der Ambivalenz  $c$  vs.  $c'$  der Abbildungspunkt des Kratermittelpunktes nicht bekannt ist. Berechnet man die Kamerapose ersatzweise mit dem Mittelpunkt der Ellipse oder rät eine der beiden Möglichkeiten, erzeugt man einen zusätzlichen Fehler in der berechneten Kamerapose.

## 3.6 Freie Oberflächenmodelle

Das im letzten Abschnitt erläuterte Kugelmodell für die Mondoberfläche ist nur realistisch bei großen Flughöhen, die von der Größenordnung des Mondradius selbst sind. Insbesondere nämlich nur solange, wie die Horizont- bzw. Oberflächenkrümmung die Geländedynamik deutlich dominiert. Liegen zu erkennende Krater auf dynamischem Gelände, erzeugt die variable Neigung des Untergrundes direkt Abweichungen in den Ellipsenparametern der abgebildeten Kraterränder. Die rekonstruierte Oberflächennormale (vgl. Abschnitt 3.3) repräsentiert dann nicht die im Modell gesuchte Referenzkugel, sondern die lokale Geländeneigung. Die Lösung von (3.31) hat in diesem Fall keine Bedeutung mehr. In Abb. 3.10 ist dieser Fall schematisch dargestellt: Es werden zwei Krater korrekt erkannt, aufgrund von Oberflächendynamik liegen die rekonstruierten Kraterränder aber koplanar, weswegen ihre Normalen  $n_1$  und  $n_2$  parallel sind. Das Modell (3.31) hätte in diesem Fall keine Lösung, da die einzige sinnvolle Lösung die eingezeichnete Ebene („Rekonstruierte Ober-

### 3.6.1 Darstellung der Oberflächenpunktmenge

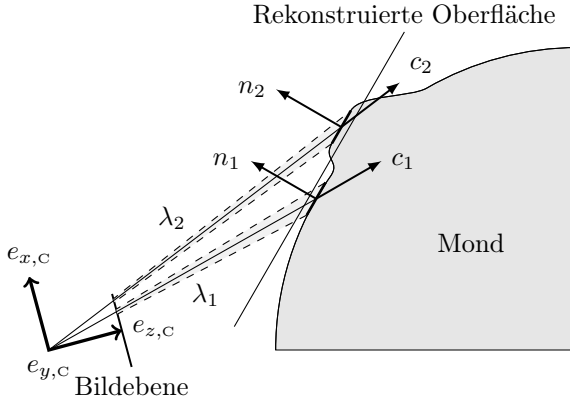


Abbildung 3.10: Oberflächendynamik verhindert Nadir-Rekonstruktion

fläche“) wäre, die aber zum als fest angenommenen Radius in Widerspruch<sup>3</sup> steht.

Man könnte das Kugelmodell um weitere Parameter erweitern, um etwa über Kugelflächenfunktionen (Kugelharmonischen) höherer Ordnung die Geländedynamik auf den bei typischen Landemissionen vorkommenden Skalen abzubilden. Dieser Ansatz würde aber weiterhin ein globales Modell aus einem in Relation winzigen lokalen Mondausschnitt schätzen und wäre in dem Sinne hoch-unterbestimmt. Genau das ist bei den „freien“ nicht-parametrischen Modellen nicht der Fall, denn diese sind ihrer Natur nach stückweise definiert und eignen sich damit besser für die lokale Oberflächenschätzung.

### 3.6.1 Darstellung der Oberflächenpunktmenge

Definition 3.2 beschreibt die formalen Anforderungen an die gesuchte Oberfläche  $S_{\kappa}^{\lambda} \subset \mathbb{R}^3$ : Sie soll minimal glatt ( $\mathcal{C}^1$ ) sein, festgemacht in Ort und Ableitung an den konfigurierten Zentren- und Normalenvektoren  $\mathbf{R}^{[\kappa]}$  und relativ in der Tiefe skaliert durch  $\lambda$ . Diese Anforderungen machen jedoch noch keinerlei Aussage darüber, wie genau die gesuchte Oberfläche parametrisiert werden kann bzw. als Funktion darzustellen ist. In diesem Abschnitt soll eine solche konsistente Darstellung hergeleitet und beschrieben werden.

<sup>3</sup>Die Ebene ist die beste Lösung, impliziert aber als Radius  $r = \infty$ , im Widerspruch zur Annahme  $r \approx 1713100\text{m}$

Viele zweidimensionale Oberflächen in dreidimensionalen Räumen kann man über zwei Variablen parametrisieren, das Finden einer solchen Parametrisierung ist bei allen Flächen, die nicht zu den geometrischen Grundkörpern (Fläche, Zylinder, Kugel, etc.) gehören, allerdings nicht trivial. Ausgehend von der Nullstellenmenge einer Funktion von drei Variablen, kann man sie zunächst implizit beschreiben:

$$F_{\text{crt}}^{\kappa, \lambda} : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad (x, y, z)^T \mapsto F_{\text{crt}}^{\kappa, \lambda}(x, y, z), \quad \kappa \text{ fest.} \quad (3.35)$$

$F$  bildet drei Variablen  $x$ ,  $y$  und  $z$  (daher der Index „crt“ für kartesisch) auf einen Skalar  $F_{\text{crt}}^{\kappa, \lambda}(x, y, z)$  ab, beschreibt also ein dreidimensionales Skalarfeld. Der Skalarsatz  $\lambda$ , der ja die Entfernungen der Oberfläche in den Richtungen  $\hat{c}_i$  vom Kameraursprung aus beschreibt, ist zunächst nicht näher bestimmt. Er ist aber als Parameter zu verstehen und nicht etwa Variable der Funktion. Die gesuchte Oberfläche ist in dieser Darstellung die Nullmenge von  $F_{\text{crt}}^{\kappa, \lambda}$ :

$$(x, y, z) \in S_{\kappa}^{\lambda} \iff F_{\text{crt}}^{\kappa, \lambda}(x, y, z) = 0. \quad (3.36)$$

Auch wenn nur diese implizite Definition der Oberfläche gegeben ist, lässt sich zumindest ihre Existenz allein aus den vorliegenden Daten  $\mathbf{R}^{[\kappa]}$  ableiten. Die gewählte implizite Darstellung der Oberfläche über die Gleichungen (3.35) und (3.36), zusammen mit der stetigen Differenzierbarkeit und der Forderung, die  $\hat{n}_i$  seien die Normalen an der Oberfläche am Punkt  $\lambda_i \hat{c}_i$  aus Def. 3.2, ist nämlich äquivalent zur Forderung, dass die  $\hat{n}_i$  mit dem JACOBI-Feld des Skalarfeldes  $F_{\text{crt}}^{\kappa, \lambda}$  an dieser Stelle übereinstimmen:

$$J[F_{\text{crt}}^{\kappa, \lambda}](\lambda_i \hat{c}_i) = \hat{n}_i, \quad i = 1, \dots, k. \quad (3.37)$$

Diese Bedingung setzt das unbekannte Skalarfeld  $F_{\text{crt}}^{\kappa, \lambda}$  in Zusammenhang mit den bekannten Größen  $\hat{c}_i$  und  $\hat{n}_i$ , sie wird daher im Folgenden als **Interpolationsbedingung** bzw. **-eigenschaft** bezeichnet werden.

Wie genau kann die Nullstellenfläche  $S_{\kappa}^{\lambda}$  gefunden werden? Das Skalarfeld  $F_{\text{crt}}^{\kappa, \lambda}$  selbst ist auch nicht explizit bekannt oder existiert möglicherweise nicht einmal als geschlossene Funktion und kann daher nicht ausgewertet werden. Es liegt lediglich die Menge der rekonstruierten Zentren- und Normalenvektoren  $\mathbf{R}$  vor und der Zusammenhang (3.37), in dem sie mit der Ableitung des Skalarfeldes steht, dessen Nullmenge die gesuchte Oberfläche ist.

Die Kraterzentrumsrichtungen und die dortigen Normalen beruhen auf der tatsächlichen Kraterverteilung in den Bilddaten und sind daher weder regelmäßig verteilt, noch liegen sie annähernd dicht im Bildbereich. Um ein

### 3.6.2 Existenz der expliziten Oberflächenfunktion

möglichst glattes Feld  $F_{\text{crt}}^{\kappa, \lambda}$  zu erhalten und gleichzeitig so wenig einschränkende Annahmen wie möglich zu machen, bietet sich ein interpolatorischer Ansatz über die Minimalvoraussetzungen aus Definition 3.2 an.

### 3.6.2 Existenz der expliziten Oberflächenfunktion

Es wird im Folgenden immer ohne weitere Erwähnung angenommen, dass die zu rekonstruierende Oberfläche beobachtbar ist. Für globale Rekonstruierbarkeit dürfen insbesondere keine Verdeckungen vorliegen (z.B. durch Berge oder Hänge), da zwischen den Messpunkten (Kratern) das Interpolations-Modell eine stetige Fortsetzbarkeit der Oberfläche voraussetzt (formal muss die echte beobachtete Oberfläche bijektiv in kamerazentrische Kugelkoordinaten abbildbar sein).

Für eine Interpolation ist es immer sinnvoll, die Dimension des Problems soweit wie möglich zu reduzieren. Die gesuchte Oberfläche ist Nullmenge eines Skalarfeldes  $F_{\text{crt}}^{\kappa, \lambda}$  und per Konstruktion zweidimensional (also Unterraum des  $\mathbb{R}^3$  und ohne Volumen). Es gilt, aus der *impliziten* Darstellung  $F_{\text{crt}}^{\kappa, \lambda}(x, y, z) = 0$  eine *explizite* Darstellung der Form  $z(x, y)$  abzuleiten, die einem Punkt  $(x, y)^T$ , etwa in Bildrahmenkoordinaten, einen Tiefen- bzw. Entfernungswert  $z(x, y)$  in Kamerakoordinaten zuordnet.

Das ist nicht notwendigerweise überhaupt möglich. Im Zweidimensionalen ist dieser Fall illustriert in Abb. 3.11. Dort lässt sich erkennen, dass für die Projektion aller auf einer ebenen Oberfläche erkannten Kraterzentren die kamerafesten  $y$ -Werte auf den selben Punkt fallen. Man erkennt sofort, dass

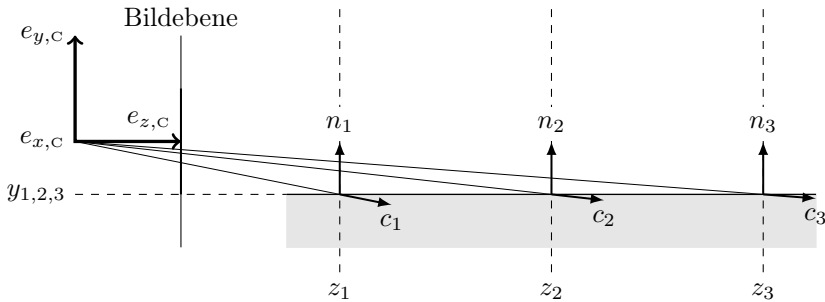


Abbildung 3.11: Stellen ohne kartesische Umkehrbarkeit

die Oberflächenfunktion an mindestens den drei markierten Stellen nicht injektiv ist: Sie besitzt für dieses „ $y$ “ ja bereits drei verschiedene „ $z$ “, ist also

nicht eindeutig umkehrbar. Dies ist auch nicht weiter überraschend, da in Abschnitt 3.3 die Oberflächennormalen aus einer *perspektivischen* Projektion von Bildpunkten konstruiert wurden. Die kartesische Darstellung „ $z(x, y)$ “ impliziert dagegen eine *orthogonale* Projektion.

Eine Lösung des Problems kann daher in einem Wechsel des Koordinatensystems in die „natürlichen“ perspektivischen Koordinaten liegen, den sphärischen oder Kugelkoordinaten.

**Definition 3.3 (Kugelkoordinaten)** Sei  $\nu \in [0, \pi]$  der Polarwinkel,  $\varphi \in [-\pi, \pi]$  der Azimuthwinkel und  $r \geq 0$  der Radius, dann beschreibt die Abbildung

$$T_{\Phi}^C : [0, \pi] \times [-\pi, \pi] \times \mathbb{R}_0^+ \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} \varphi \\ \nu \\ r \end{pmatrix} \mapsto \begin{pmatrix} r \sin \nu \cos \varphi \\ r \sin \nu \sin \varphi \\ r \cos \nu \end{pmatrix} \quad (3.38)$$

die Transformation von Kugelkoordinaten in die kartesischen Koordinaten.

Der Zusammenhang mit den weiteren bereits eingeführten Koordinatensystemen, dem Bild-, Bildrahmen-, und dem kartesischen Kamerakoordinatensystem, ist in Abbildung 3.12 auf Seite 66 zusammengefasst dargestellt<sup>4</sup>.

Illustriert wird dieser Zusammenhang der Koordinatensysteme in Abbildung 3.13 auf Seite 67 anhand eines Beispiels mit einer planen Oberfläche, auf der zwei Krater erkannt wurden. Die rekonstruierten räumlichen Zentrumsrichtungen  $c_i$  und  $n_i$  und ihre Abbildung auf die Bildebene sind dort zu erkennen.

Mit diesen Definitionen lässt sich jedes kartesische Skalarfeld  $F_{\text{crt}}^{\kappa, \lambda}(x, y, z)$  aus (3.35) in diesen sphärischen Koordinaten ausdrücken:

$$\begin{aligned} F_{\text{sph}}^{\kappa, \lambda}(\nu, \varphi, r) &= (F_{\text{crt}}^{\kappa, \lambda} \circ T_{\Phi}^C)(\nu, \varphi, r) \\ &= F_{\text{crt}}^{\kappa, \lambda}(r \sin \nu \cos \varphi, r \sin \nu \sin \varphi, r \cos \nu), \end{aligned} \quad (3.39)$$

wobei „ $\circ$ “ die Verkettung von Funktionen anzeigt (sprich „ $F_{\text{crt}}^{\kappa, \lambda}$  nach  $T_{\Phi}^C$  von  $\nu, \varphi, r$ “). Hiermit sei die folgende Behauptung motiviert:

<sup>4</sup>In der Abbildung sind auch an den sphärischen Koordinaten die Indizes, die das jeweilige Koordinatensystem anzeigen, immer vorhanden. Im Folgenden wird dieser Index bei Unmissverständlichkeit immer weggelassen.

**Behauptung 3.1 (Lokale Existenz der expliziten Oberflächenfunktion)**

Es seien  $F_{\text{crt}}^{\kappa, \lambda}$  die kartesische implizite Darstellung einer Oberfläche gemäß Def. 3.2 mit der Interpolationseigenschaft aus Gleichung (3.36), und  $F_{\text{sph}}^{\kappa, \lambda}$  ihre Entsprechung in Kugelkoordinaten gemäß Gleichung (3.39). Dann existiert für alle  $0 < \lambda < \infty$  (komponentenweise) an allen Stellen

$$\begin{pmatrix} \lambda_i \sin \nu_i \cos \varphi_i \\ \lambda_i \sin \nu_i \sin \varphi_i \\ \lambda_i \cos \nu_i \end{pmatrix} = \lambda_i \hat{c}_i \quad (3.40)$$

für  $\hat{c}_i$  aus  $\mathbf{R}^{[\kappa]}$  eine explizite Darstellung  $r(\nu, \varphi)$ , für die gilt

$$F_{\text{sph}}^{\kappa, \lambda}(\nu_i, \varphi_i, r(\nu_i, \varphi_i)) = 0, \quad (3.41)$$

diese Funktion  $r : [0, \pi] \times [-\pi, \pi] \rightarrow \mathbb{R}_0^+$  beschreibt die gesuchte Nullstellenmenge  $S_{\kappa}^{\lambda}$  des skalaren Feldes  $F_{\text{crt}}^{\kappa, \lambda}$  explizit in einer nichtverschwindenden Umgebung um die Stützstellen  $\lambda_i \hat{c}_i$ .

**Beweis:** Nach dem Satz von der impliziten Funktion existiert das gesuchte  $r(\nu, \varphi)$  genau dann in einer Umgebung um eine beliebige der Stützstellen  $\lambda_i \hat{c}_i$ , wenn der explizite Anteil (der radiale Anteil  $r$ ) der JACOBI-Matrix  $J[F_{\text{sph}}^{\kappa, \lambda}]$  des Skalarfeldes dort invertierbar ist.

Das Skalarfeld in Kugelkoordinaten ist nach Gleichung (3.39) über die Verkettung mit der Koordinatentransformation gegeben, seine JACOBI-Matrix folgt danach aus der Kettenregel:

$$\begin{aligned} J[F_{\text{sph}}^{\kappa, \lambda}](\nu, \varphi, r) &= J[F_{\text{crt}}^{\kappa, \lambda} \circ T_{\Phi}^{\text{c}}](\nu, \varphi, r) \\ &= J[F_{\text{crt}}^{\kappa, \lambda}](T_{\Phi}^{\text{c}}(\nu, \varphi, r)) \cdot J[T_{\Phi}^{\text{c}}](\nu, \varphi, r), \end{aligned} \quad (3.42)$$

der Punkt „ $\cdot$ “ bezeichnet die Matrixmultiplikation. Die Transformation  $T_{\Phi}^{\text{c}}$  aus Gleichung (3.38) besitzt die folgende  $3 \times 3$ -JACOBI-Matrix:

$$\begin{aligned} J[T_{\Phi}^{\text{c}}](\nu, \varphi, r) &= \begin{pmatrix} \frac{\partial T_{\Phi 1}^{\text{c}}}{\partial \nu}(\nu, \varphi, r) & \frac{\partial T_{\Phi 1}^{\text{c}}}{\partial \varphi}(\nu, \varphi, r) & \frac{\partial T_{\Phi 1}^{\text{c}}}{\partial r}(\nu, \varphi, r) \\ \frac{\partial T_{\Phi 2}^{\text{c}}}{\partial \nu}(\nu, \varphi, r) & \frac{\partial T_{\Phi 2}^{\text{c}}}{\partial \varphi}(\nu, \varphi, r) & \frac{\partial T_{\Phi 2}^{\text{c}}}{\partial r}(\nu, \varphi, r) \\ \frac{\partial T_{\Phi 3}^{\text{c}}}{\partial \nu}(\nu, \varphi, r) & \frac{\partial T_{\Phi 3}^{\text{c}}}{\partial \varphi}(\nu, \varphi, r) & \frac{\partial T_{\Phi 3}^{\text{c}}}{\partial r}(\nu, \varphi, r) \end{pmatrix} \\ &= \begin{pmatrix} r \cos \nu \cos \varphi & -r \sin \nu \sin \varphi & \sin \nu \cos \varphi \\ r \cos \nu \sin \varphi & r \sin \nu \cos \varphi & \sin \nu \sin \varphi \\ -r \sin \nu & 0 & \cos \nu \end{pmatrix}, \end{aligned} \quad (3.43)$$

wobei die Indizierung  $T_{\Phi i}^C$ ,  $i = 1, 2, 3$  die jeweilige der drei Komponenten von  $T_{\Phi}^C$  anzeigt. Die JACOBI-Matrix von  $F_{\text{crt}}^{\kappa, \lambda}$  ist die  $1 \times 3$ -Matrix

$$J [F_{\text{crt}}^{\kappa, \lambda}] (x, y, z) = \begin{pmatrix} \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial x}(x, y, z) & \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial y}(x, y, z) & \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial z}(x, y, z) \end{pmatrix} \quad (3.44)$$

Einsetzen von (3.43) und (3.44) in (3.42) liefert

$$J [F_{\text{sph}}^{\kappa, \lambda}] = \begin{pmatrix} \left( \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial x} \circ T_{\Phi}^C \right) \frac{\partial T_{\Phi 1}^C}{\partial \nu} + \left( \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial y} \circ T_{\Phi}^C \right) \frac{\partial T_{\Phi 2}^C}{\partial \nu} + \left( \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial z} \circ T_{\Phi}^C \right) \frac{\partial T_{\Phi 3}^C}{\partial \nu} \\ \left( \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial x} \circ T_{\Phi}^C \right) \frac{\partial T_{\Phi 1}^C}{\partial \varphi} + \left( \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial y} \circ T_{\Phi}^C \right) \frac{\partial T_{\Phi 2}^C}{\partial \varphi} + \left( \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial z} \circ T_{\Phi}^C \right) \frac{\partial T_{\Phi 3}^C}{\partial \varphi} \\ \left( \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial x} \circ T_{\Phi}^C \right) \frac{\partial T_{\Phi 1}^C}{\partial r} + \left( \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial y} \circ T_{\Phi}^C \right) \frac{\partial T_{\Phi 2}^C}{\partial r} + \left( \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial z} \circ T_{\Phi}^C \right) \frac{\partial T_{\Phi 3}^C}{\partial r} \end{pmatrix}^T \quad (3.45)$$

Der Kürze halber steht (3.45) in Operatorform, die Argumente  $(\nu, \varphi, r)$  sind wie in (3.43) und (3.44) auf jeden Faktor jedes Summanden jeder Komponente anzuwenden. Man beachte auch, dass (3.45) transponiert steht, also ein Zeilenvektor, ist.

Der „ $\partial r$ -Anteil“ (der explizite Anteil, letzte Zeile von (3.45)) ist also ausgeschrieben

$$\begin{aligned} \frac{\partial F_{\text{sph}}^{\kappa, \lambda}}{\partial r}(\nu, \varphi, r) &= \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial x}(T_{\Phi}^C(\nu, \varphi, r)) \sin \nu \cos \varphi \\ &+ \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial y}(T_{\Phi}^C(\nu, \varphi, r)) \sin \nu \sin \varphi \\ &+ \frac{\partial F_{\text{crt}}^{\kappa, \lambda}}{\partial z}(T_{\Phi}^C(\nu, \varphi, r)) \cos \nu. \end{aligned} \quad (3.46)$$

Gleichung (3.46) ist Skalar, also als  $1 \times 1$ -Matrix genau dann invertierbar, wenn sie nicht 0 ist. Da die Auswertung an den Stützstellen  $\lambda_i \hat{c}_i$  stattfindet und

$$\lambda_i \hat{c}_i = T_{\Phi}^C(\nu_i, \varphi_i, \lambda_i) = (\lambda_i \sin \nu_i \cos \varphi_i \quad \lambda_i \sin \nu_i \sin \varphi_i \quad \lambda_i \cos \nu_i)^T \quad (3.47)$$

ist, wird (3.46) mit der Interpolationseigenschaft aus Gleichung (3.37) zu

$$\begin{aligned} \frac{\partial F_{\text{sph}}^{\kappa, \lambda}}{\partial r}(\nu_i, \varphi_i) &\stackrel{(3.37)}{=} \hat{n}_{i,1} \sin \nu_i \cos \varphi_i + \hat{n}_{i,2} \sin \nu_i \sin \varphi_i + \hat{n}_{i,3} \cos \nu_i \\ &\stackrel{(3.47)}{=} \hat{n}_{i,1} \hat{c}_{i,1} + \hat{n}_{i,2} \hat{c}_{i,2} + \hat{n}_{i,3} \hat{c}_{i,3} \\ &= \hat{c}_i^T \hat{n}_i. \end{aligned} \quad (3.48)$$

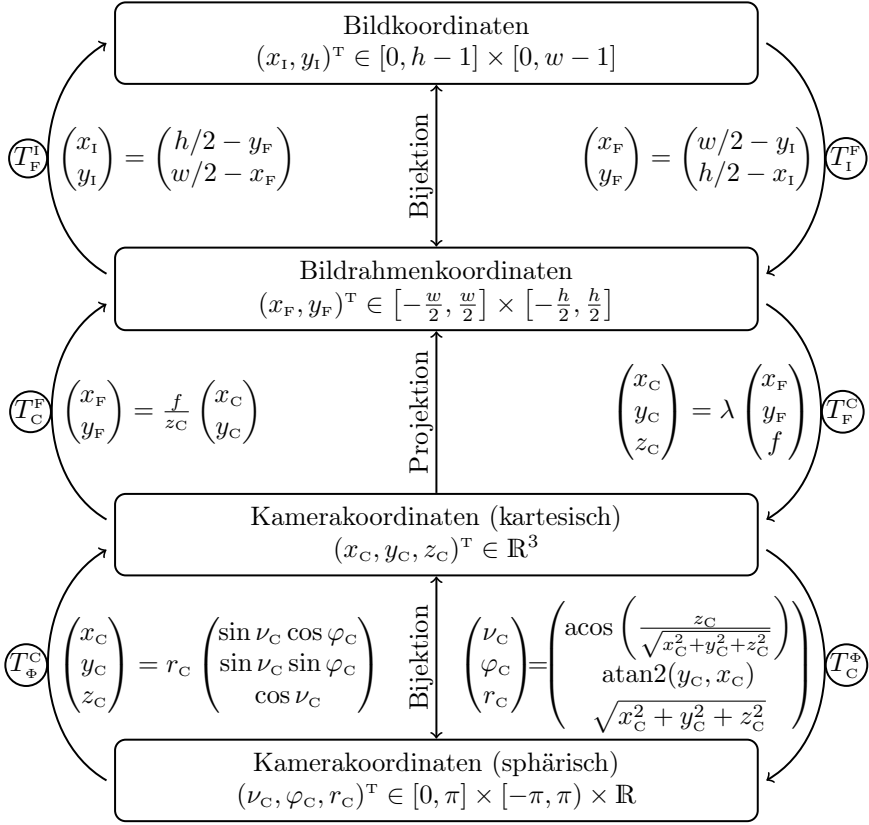


Abbildung 3.12: Von Bild- nach sphärischen Kamerakoordinaten

$\lambda_i$  ist per Definition 3.2 größer 0 und kleiner  $\infty$ , die Gleichung (3.48) kann also nur dann 0 werden, wenn die Vektoren des  $i$ -ten Zentren-Normalenpaares aus der Konfiguration  $\mathbf{R}^{[\kappa]}$  orthogonal aufeinander stehen. Das ist ausgeschlossen, da es voraussetzen würde, dass  $\hat{c}_i$  und  $\hat{n}_i$  aus einer degenerierten Ellipse (eine Linie) rekonstruiert worden wäre. Solche Ellipsen können nicht gültiges Ergebnis der Kratererkennung sein, die Behauptung ist also wahr.  $\square$

Dies sichert die Existenz einer expliziten Darstellung einer gegebenen Fläche  $S_\kappa^\lambda \in \mathbb{R}^3$  durch eine sphärische Funktion  $r(\nu, \varphi)$  in einer nichtverschwindenden Umgebung um alle Punkte  $\lambda_i \hat{c}_i$ . Als nächstes muss untersucht



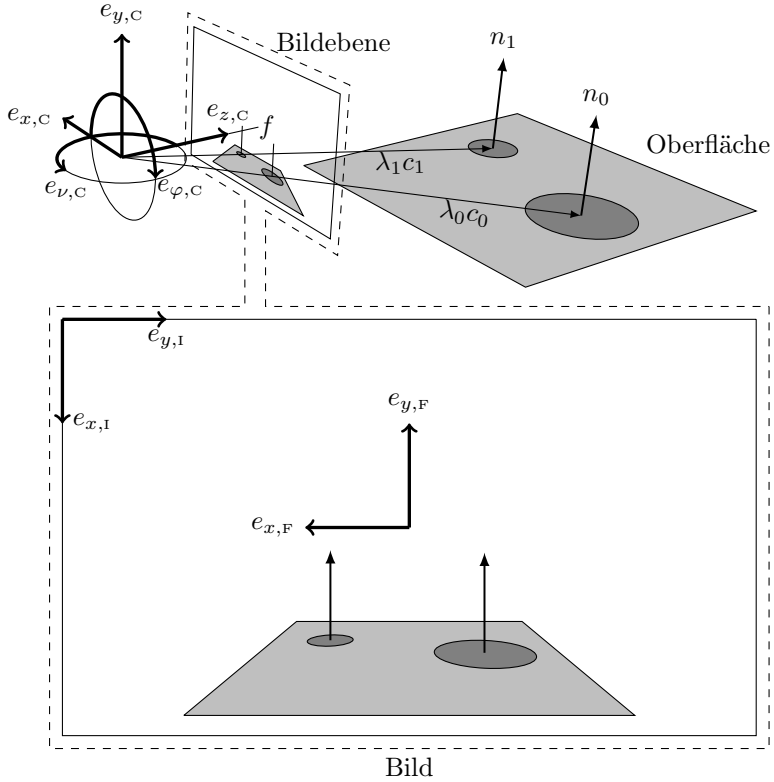


Abbildung 3.13: Perspektivische Projektion von räumlichen Kratern in eine Bildebene und Zusammenhang mit sphärischen Koordinaten.

werden, ob und wie sich speziell aus den Daten  $\mathbf{R}$  eine solche Funktion rekonstruieren lässt.

**Behauptung 3.2 (Rekonstruierbarkeit aus den Daten)** Zu einer existenten aber unbekannten expliziten Darstellung  $r(\nu, \varphi)$  einer gegebenen impliziten Oberfläche  $S_{\kappa}^{\lambda}$  gibt es eine feste Konstante  $\Lambda$  und (mindestens) eine aus den Daten rekonstruierbare sphärische Funktion  $\hat{r}(\nu, \varphi)$ , für die gilt

$$r(\nu_i, \varphi_i) = \hat{r}(\nu_i, \varphi_i) \cdot \Lambda \quad (3.49)$$

### 3.6.2 Existenz der expliziten Oberflächenfunktion

an allen Stützstellen

$$\begin{pmatrix} \sin \nu_i \cos \varphi_i \\ \sin \nu_i \sin \varphi_i \\ \cos \nu_i \end{pmatrix} = \hat{c}_i \quad (3.50)$$

einer fest konfigurierten  $\mathbf{R}^{[\kappa]}$  Zentren- und Normalenvektoren-Menge. Für dieses  $\hat{r}(\nu, \varphi)$  lassen sich die Funktionswerte an den Stellen  $(\nu_i, \varphi_i)$  aus den Daten rekonstruieren.

**Beweis:** An allen Stützstellen  $\lambda_i \hat{c}_i$  ist  $r(\nu, \varphi)$  differenzierbar nach  $\nu$  und  $\varphi$ , weil  $F_{\text{sph}}^{\kappa, \lambda}$  es ist. Aus der verallgemeinerten Kettenregel ergibt sich der Gradient von  $r$ :

$$\begin{aligned} \nabla r(\nu, \varphi) &= \begin{pmatrix} \frac{\partial r}{\partial \nu}(\nu, \varphi) & \frac{\partial r}{\partial \varphi}(\nu, \varphi) \end{pmatrix} \\ &= - \begin{pmatrix} \frac{J_1[F_{\text{sph}}^{\kappa, \lambda}](\nu, \varphi, r(\nu, \varphi))}{J_3[F_{\text{sph}}^{\kappa, \lambda}](\nu, \varphi, r(\nu, \varphi))} & \frac{J_2[F_{\text{sph}}^{\kappa, \lambda}](\nu, \varphi, r(\nu, \varphi))}{J_3[F_{\text{sph}}^{\kappa, \lambda}](\nu, \varphi, r(\nu, \varphi))} \end{pmatrix} \end{aligned} \quad (3.51)$$

(Die Indizes an der JACOBI-Matrix indizieren ihre Komponente.) Der  $J[T_{\Phi}^c]$ -Anteil der dritten Komponente  $J_3[F_{\text{sph}}^{\kappa, \lambda}]$  der JACOBI-Matrix enthält bereits kein  $r(\nu, \varphi)$  mehr (die rechte Spalte der Matrix in Gleichung (3.43)). Aus den ersten beiden Komponenten lässt es sich herausdividieren, so, dass es überall nur noch im kartesischen JACOBI-Anteil auftritt. Man erhält

$$\frac{\nabla r(\nu, \varphi)}{r(\nu, \varphi)} = - \begin{pmatrix} \frac{(J[F_{\text{crt}}^{\kappa, \lambda}] \circ T_{\Phi}^c)(\nu, \varphi, r(\nu, \varphi)) \cdot \begin{pmatrix} \cos \nu \cos \varphi \\ \cos \nu \sin \varphi \\ -\sin \nu \end{pmatrix}}{(J[F_{\text{crt}}^{\kappa, \lambda}] \circ T_{\Phi}^c)(\nu, \varphi, r(\nu, \varphi)) \cdot \begin{pmatrix} \sin \nu \cos \varphi \\ \sin \nu \sin \varphi \\ \cos \nu \end{pmatrix}} \\ \frac{(J[F_{\text{crt}}^{\kappa, \lambda}] \circ T_{\Phi}^c)(\nu, \varphi, r(\nu, \varphi)) \cdot \begin{pmatrix} -\sin \nu \sin \varphi \\ \sin \nu \cos \varphi \\ 0 \end{pmatrix}}{(J[F_{\text{crt}}^{\kappa, \lambda}] \circ T_{\Phi}^c)(\nu, \varphi, r(\nu, \varphi)) \cdot \begin{pmatrix} \sin \nu \cos \varphi \\ \sin \nu \sin \varphi \\ \cos \nu \end{pmatrix}} \end{pmatrix}^T. \quad (3.52)$$

Setzt man jetzt die Stützstellen  $\lambda_i \hat{c}_i$  ein, wird Gleichung (3.52) wegen der

Interpolationsbedingung (3.37) und Gleichung (3.48) zu

$$\frac{\nabla r(\nu_i, \varphi_i)}{r(\nu_i, \varphi_i)} = -\frac{\hat{n}_i^T}{\hat{c}_i^T \hat{n}_i} \begin{pmatrix} \cos \nu_i \cos \varphi_i & -\sin \nu_i \sin \varphi_i \\ \cos \nu_i \sin \varphi_i & \sin \nu_i \cos \varphi_i \\ -\sin \nu_i & 0 \end{pmatrix}. \quad (3.53)$$

An den Stützstellen ist die rechte Seite frei von Unbekannten und lässt sich aus den Daten berechnen. Wähle jetzt

$$\hat{r}(\nu, \varphi) := \exp \left( \int \frac{\nabla r(\nu, \varphi)}{r(\nu, \varphi)} \right) \quad (3.54)$$

und

$$\Lambda := \frac{\min_{\hat{c}_i} \{r(\nu, \varphi)\}}{\min_{\hat{c}_i} \{\hat{r}(\nu, \varphi)\}}. \quad (3.55)$$

Der Ausdruck  $\nabla r(\nu, \varphi)/r(\nu, \varphi)$  besitzt Stammfunktionen

$$\int \frac{\nabla r(\nu, \varphi)}{r(\nu, \varphi)} = \ln(r(\nu, \varphi)) + \Gamma \quad (3.56)$$

für konstante  $\Gamma \in \mathbb{R}$  (man überprüfe (3.56) anhand der Kettenregel). Mit (3.54) und (3.55) löst sich Gleichung (3.49) über

$$\begin{aligned} \hat{r}(\nu, \varphi) \cdot \Lambda &= \exp \left( \int \frac{\nabla r(\nu, \varphi)}{r(\nu, \varphi)} \right) \cdot \frac{\min_{\hat{c}_i} \{r(\nu, \varphi)\}}{\exp \left( \min_{\hat{c}_i} \left\{ \int \frac{\nabla r(\nu, \varphi)}{r(\nu, \varphi)} \right\} \right)} \\ &= \exp \left( \int \frac{\nabla r(\nu, \varphi)}{r(\nu, \varphi)} \right) \\ &\quad \cdot \exp \left( \ln \left( \min_{\hat{c}_i} \{r(\nu, \varphi)\} \right) - \min_{\hat{c}_i} \left\{ \int \frac{\nabla r(\nu, \varphi)}{r(\nu, \varphi)} \right\} \right) \\ &= \exp \left( \ln(r(\nu, \varphi)) + \Gamma + \ln(\min_{\hat{c}_i} r(\nu, \varphi)) \right. \\ &\quad \left. - \min_{\hat{c}_i} \{\ln(r(\nu, \varphi)) + \Gamma\} \right) \\ &= \exp(\ln(r(\nu, \varphi))) \\ &= r(\nu, \varphi). \end{aligned} \quad (3.57)$$

(Sowohl der natürliche Logarithmus „ln()“, als auch seine Umkehrfunktion, die Exponentialfunktion „exp()“ sind monoton, lassen sich daher mit der

### 3.6.2 Existenz der expliziten Oberflächenfunktion

---

Minimumbildung vertauschen.)

□

In Zusammenfassung des bis zu diesem Punkt der Arbeit Vorausgesetzten und Gezeigten:

- Zu einer vorausgesetzten „echten“ Oberfläche  $S_{\kappa}^{\lambda}$  ist nur bekannt, dass sie die Eigenschaften aus Def. 3.2 erfüllt, dass sie keine Verdeckungen bzw. Hinterschnidungen besitzt und dass ihre Oberflächennormalen an Punkten unbekannter Entfernung auf bekannten Sichtstrahlen vorliegen.
- Der Beweis von Behauptung 3.1 liefert die Sicherheit, dass die 3D-Zentren- und Normalendaten mit einer expliziten 2D-Darstellung  $r(\nu, \varphi)$  kompatibel sind (lokale Invertierbarkeit bzgl. der Koordinatentransformation).
- Die Fortsetzbarkeit auf das gesamte abgedeckte Gebiet folgt aus der Glattheitsanforderung an  $S_{\kappa}^{\lambda}$  und aus der Voraussetzung, dass die Oberfläche keine Hinterschnidungen bezüglich der Kugelkoordinaten  $\nu, \varphi$  haben durfte.
- Behauptung 3.2 sagt, dass sich eine Funktion  $\hat{r}(\nu, \varphi)$  finden lässt, die aus den Daten  $\mathbf{R}^{[\kappa]}$  rekonstruierbar ist und die an den Datenpunkten bis auf Skalierung dem unbekannten  $r(\nu, \varphi)$  entspricht.
- $\hat{r}(\nu, \varphi)$  ist eine Funktion, die man erhält, wenn man den skalierten Gradienten  $\nabla r(\nu_i, \varphi_i)/r(\nu_i, \varphi_i)$  auswertet (Gleichung (3.53)), gemäß den Glattheitsanforderungen auf alle  $(\nu, \varphi)$  fortsetzt, integriert und exponenziert (Gleichung (3.54)).

Ganz konkret wurde das Problem der Schätzung einer völlig unbekannten dreidimensionalen Oberfläche aus einigen Messungen ihrer Oberflächennormalen reduziert auf ein klassisches zweidimensionales Interpolationsproblem. Die Güte der Approximation durch ein dadurch gefundenes  $\hat{r}(\nu, \varphi)$  hängt davon ab, wie getreu die tatsächliche Mondoberfläche, auf der die Messdaten beruhen, den Anforderungen der Def. 3.2 entspricht und welches Modell zur Interpolation angesetzt wird.

Um das praktische Vorgehen zu verdeutlichen, ist der in Abbildung 3.11 zur Motivation des Wechsels auf die sphärischen Koordinaten illustrierte Fall beispielhaft in Bsp. A.2 im Anhang A durchgerechnet. Als Interpolationsmethode wird dort schlicht ein interpolierendes Polynom über die Werte der Ableitung gebildet und integriert.

Die Oberflächenrekonstruktion wird mit dem Ziel betrieben, zur Korrespondenzbildung mit dem bekannten Kraterkatalog die tatsächlichen Größenverhältnisse erkannter Krater, Entfernungen zwischen ihnen und Innenwinkel von aus ihnen gebildeten Konstellationen zu bestimmen. Alle diese gesuchten Größen sind *relativ*, die Rekonstruktion der Oberfläche bis auf die globale Skalierungskonstante  $\Lambda$  ist also für diese Zwecke ausreichend.

Veranschaulichung: Seien  $\rho_1$  und  $\rho_2$  die tatsächlichen physischen Radien zweier benachbarter Krater, die im Bild erkannt wurden. Die Kratermittelpunkte liegen bei  $\lambda_1 \hat{c}_1$  und  $\lambda_2 \hat{c}_2$  im Kamera-Koordinatensystem. Bezeichne  $\bar{\rho}_1$  und  $\bar{\rho}_2$  den „scheinbaren“ Radius<sup>5</sup> der Krater im Einheitsabstand, also für  $\lambda_{1,2} = 1$ . Da der tatsächliche Abstand zur wahren Oberfläche im Kamerakoordinatensystem immer unbekannt ist, sind stets nur  $\bar{\rho}_1$  und  $\bar{\rho}_2$  bekannt.

Sei weiterhin mit  $r(\nu, \varphi)$  die explizite Darstellung der wahren Oberfläche bezeichnet, so dass

$$\rho_i = \bar{\rho}_i \cdot r(\nu_i, \varphi_i) \text{ für } \begin{pmatrix} \sin \nu_i \cos \varphi_i \\ \sin \nu_i \sin \varphi_i \\ \cos \nu_i \end{pmatrix} = \hat{c}_i, \quad i = 1, 2. \quad (3.58)$$

Für das Verhältnis der Radien der beiden Krater folgt mit der rekonstruierten Oberflächenfunktion  $\hat{r}(\nu, \varphi)$  daraus

$$\frac{\rho_1}{\rho_2} = \frac{\bar{\rho}_1 \cdot r(\nu_1, \varphi_1)}{\bar{\rho}_2 \cdot r(\nu_2, \varphi_2)} = \frac{\bar{\rho}_1 \cdot \hat{r}(\nu_1, \varphi_1) \cdot \Lambda}{\bar{\rho}_2 \cdot \hat{r}(\nu_2, \varphi_2) \cdot \Lambda} = \frac{\bar{\rho}_1 \cdot \hat{r}(\nu_1, \varphi_1)}{\bar{\rho}_2 \cdot \hat{r}(\nu_2, \varphi_2)}. \quad (3.59)$$

Der Wert des Quotienten lässt sich also ohne Kenntnis der Konstante in Gleichung (3.49) ermitteln, es entsteht bis auf die unbekannte Skalierung kein relativer Fehler in der Oberflächenrekonstruktion.

## 3.7 Algorithmus zur Berechnung einer Funktion der Oberfläche

### 3.7.1 Einleitung

In den Abschnitten 3.6.1 und 3.6.2 wurde die formale Existenz und Rekonstruierbarkeit einer die Daten interpolierenden Oberfläche gezeigt. Es wird jetzt ein

---

<sup>5</sup>Derjenige Radius, den der Schnitt des projektiven Kegels der erkannten Ellipse in Kamerakoordinaten mit der Kraternormalen-Ebene am Punkt mit Abstand 1 vom Kamera-Ursprung hat. Vergleiche Abschnitt 3.3

### 3.7.2 Definitionsbereich

---

Ansatz benötigt, wie diese Rekonstruktion algorithmisch ausgeführt werden kann. Das Rechenbeispiel A.2 illustriert den Fall einer eindimensionalen Oberfläche (einer Kurve in der Ebene), für den Anwendungsfall Kraternavigation wird dagegen ein Verfahren benötigt, das eine zweidimensionale Hyperfläche im dreidimensionalen Kamera-Koordinatensystem erzeugt.

Die Anforderungen an das Ergebnis eines solchen Rekonstruktionsverfahrens wurden in Definition 3.2 beschrieben: Es soll eine hinreichend glatte Flächenfunktion gefunden werden, die an den Durchstoßpunkten der Projektionslinien der erkannten Kraterzentren orthogonal auf den rekonstruierten Kraternormalen steht.

Es liegt das im vorangegangenen Abschnitt konstruierte zweidimensionale Interpolationsproblem vor. Seine Daten sind die Projektionsgeraden der erkannten Kratermittelpunkte und die rekonstruierten Kraternormalen, zusammengefasst in der Datenmenge  $\mathbf{R}$  mit nicht näher bestimmter Konfiguration  $\kappa$ . Die gesuchte Interpolante zu den Daten  $\nabla r/r$  ist eine Funktion  $s(\nu, \varphi)$ , deren Exponential wie in Gleichung (3.56) die gesuchte Oberflächenrekonstruktion  $\hat{r}(\nu, \varphi)$  ist und in vollständiger Auswertung auf ihrem Träger die Punktmenge  $S_\kappa^\lambda$  der gesuchten Oberfläche gemäß Definition 3.2 erzeugt.

In den folgenden Unterabschnitten wird das Interpolationsproblem formal gründlich eingeführt und eine praktisch sinnvolle Methode zu seiner Lösung vorgestellt.

### 3.7.2 Definitionsbereich

Das allgemeine Lochkamera-Modell und damit der Zusammenhang zwischen den Welt- bzw. Kamera-Koordinaten und ihre perspektivische Projektion auf eine Abbildungsebene wurden in Abschnitt 3.2 erläutert. Abbildung 3.13 auf Seite 67 vereinte diese Definitionen mit der verwendeten Konvention der sphärischen Winkel-Koordinaten des Kamerasystems  $\nu$  und  $\varphi$ . Das folgende virtuelle Koordinatensystem hat keine physikalische Bedeutung, es dient bloß der Konditionierung des Interpolationsproblems.

**Definition 3.4 (Virtuelle Bildkoordinaten)** Es bezeichne  $(x, y, z)^T$  einen Punkt im kartesischen Kamerakoordinatensystem. Die Transformation

$$T_C^V : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (x, y, z)^T \mapsto T_C^\Phi(z, y, -x)^T$$

beschreibt die sphärische Darstellung der positiv um  $\pi/2$  um die Kamerafeste  $e_{y,C}$ -Achse rotierten Kamerakoordinate. Ihre Inverse wird mit  $T_V^C$  bezeichnet.

Die Behauptungen zur Existenz (Beh. 3.1) und Rekonstruierbarkeit (Beh. 3.2) der Oberflächenfunktion aus den Daten der Kratererkennung sind trivialerweise auch für die virtuellen Bildkoordinaten wahr, da nur Komponenten und Vorzeichen getauscht werden. Der Definitionsbereich der Interpolante  $s(\nu, \varphi)$  lässt sich damit wie folgt beschreiben:

**Definition 3.5 (Definitionsbereich)** Es sei  $\xi = (\nu, \varphi)^T$  eine Winkelkoordinate im oben definierten virtuellen Bildkoordinatensystem. Die Menge  $\mathbb{D} \subset \mathbb{R}^2$  aller Korrespondierenden  $\xi$  und  $(x, y, z)^T = T_V^C(\xi)$ , für die gilt

$$\left| \frac{x}{z} \right| \leq \frac{w}{2f} \quad \text{und} \quad \left| \frac{y}{z} \right| \leq \frac{h}{2f},$$

mit  $w$ ,  $h$  und  $f$  die Breite, Höhe und Brennweite des Kamerabildes (in Pixeln), wird der Definitionsbereich der Interpolante  $s(\nu, \varphi) = s(\xi)$  genannt.

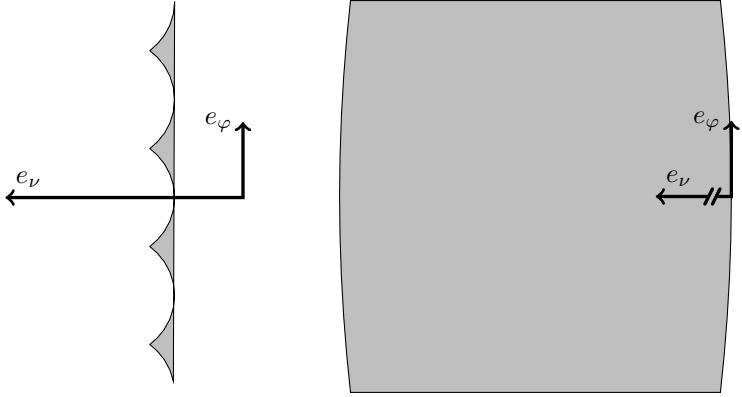
Anschaulich bedeutet diese Definition, dass  $s$  (und damit  $\hat{r}$ ) auf demjenigen abgeschlossenen Unterraum der virtuellen Bildkoordinaten definiert ist, dessen Transformation in die Kamerakoordinaten und Projektion auf die Bildebene nicht das physische Bildrechteck verlässt. Selbst muss (und kann)  $\mathbb{D}$  nicht rechteckig sein in einem kartesischen Sinn, ist aber aufgrund der Konstruktion der virtuellen Bildkoordinaten nach Def. 3.4 zumindest konvex. Dies ist in den Abbildungen 3.14a und 3.14b sichtbar: Auf der linken Seite ist der direkt aus dem physischen Kamera-Koordinatensystem in den Winkelbereich transformierte Bildrand zu sehen, auf der rechten Seite wurde der Bildrand in das virtuelle Bildkoordinatensystem transformiert. In den oben definierten virtuellen Bildkoordinaten lassen sich die im Bild erkannten Krater verorten. Folgende Definition beschreibt eine „mittlere“ Koordinate im virtuellen Bild, die die unklar konfigurierten Erkennungsdaten  $\mathbf{R}$  aus Definition 3.1 auf einen niederdimensionalen Satz Punkte reduziert:

**Definition 3.6 (Mittlere virtuelle Kraterkoordinaten)** Sei  $\mathbf{R}$  die Menge der unkonfigurierten Zentren- und Normalenvektoren gemäß Definition 3.1. Die Menge  $\boldsymbol{\xi} = \{\xi_1, \dots, \xi_{N_p}\}$  bestehend aus

$$\xi_i = (\nu_i, \varphi_i)^T = T_C^V \left( n_i + \frac{n'_i - n_i}{2} \right)$$

mit  $n_i$  und  $n'_i$  aus  $\mathbf{r}_i$  und  $\mathbf{r}'_i$ , beide in  $R_i \in \mathbf{R}$ , wird die Menge der virtuellen

### 3.7.3 Bivariater Spline-Ansatz



(a) Bildrand direkt in sphärische Koordinaten transformiert

(b) Bildrand transformiert in virtuelles Bildkoordinatensystem

Abbildung 3.14: Darstellung des Rechtecks des Bildrandes nach Transformation in sphärische und virtuelle Bildkoordinaten

Kraterkoordinaten genannt. „ $T_C^v$ “ ist dabei wieder die Transformation von kartesischen Kamerakordinaten in die virtuellen Bildkoordinaten (vgl. Def. 3.4).

Die Punktmenge  $\xi$  ist indizierbar über einen geordneten Indexvektor  $\mathcal{I} := (i_1, i_2, \dots, i_N)$  der Länge  $N$  für  $i_1 > 0$ ,  $i_k > i_{k-1}$  und  $i_N < N_p$  auf die folgende Weise:

$$\xi_{\mathcal{I}} := (\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_N})$$

Durch Konstruktion dieser mittleren Punkte liefert die Definition einen Satz virtuelle Bildkoordinaten für die Datenmenge der erkannten Krater, die konstant bei Umkonfiguration der zugrundeliegenden Daten ist.

### 3.7.3 Bivariater Spline-Ansatz

Der Definitionsbereich der gesuchten Interpolante  $s(\xi)$  ist per Def. 3.5 zweidimensional, reell, konvex und enthält wegen ihrer Konstruktion in Def. 3.6 alle Stützstellen  $\xi_i = (\nu_i, \varphi_i)$  des Interpolationsproblems.

Da die zu rekonstruierende Oberfläche glatt sein soll, bieten sich niedergradige Polynome als Modellfunktionen an. Liegen  $N_p$  Stützstellen vor, würde



man aber bereits im Eindimensionalen zur Interpolation ein Polynom vom Grad  $N_p - 1$  oder höher benötigen. Dieses besitzt dann aber auch genauso viele Extremstellen und neigt zum Aufschwingen, wenn es zur Interpolation von Funktionen benutzt wird, die selbst keine Polynome sind (das sog. RUNGE-Phänomen).

Als Alternative dazu ist die Spline-Interpolation weitestgehend unsensibel gegenüber diesem Phänomen. Sie umfasst die Konstruktion einer *stückweise* niedergradigen Polynom-Interpolante, die an den Stückgrenzen gewissen Glattheitsbedingungen genügt. Auf irregulären zweidimensionalen Gittern, wie sie im Anwendungsfall vorliegen (die erkannten Krater sind ja nicht regelmäßig im Bild verteilt), kommen Splines auf sogenannten Triangulationen zum Einsatz.

Bei einer Spline-Interpolation sind immer mindestens zwei Parameter zu wählen: Der Grad der Teilpolynome des Splines, sowie die Glattheit an den Übergängen eines der Teilpolynome zum nächsten. Da die vorliegende Arbeit einen Algorithmus zum Thema hat, der im Hinblick auf einen Einsatz auf stark leistungsbeschränkter weltraumtauglicher Hardware entwickelt wurde, werden diese Parameter so minimal wie praktikabel gewählt:

- Der Spline wird aus *kubischen* Polynomen (also vom Grad 3) konstruiert werden.
- Der Spline soll überall mindestens einmal stetig differenzierbar sein.

Da die Kraternormalen direkt äquivalent zu den Gradienten der Oberflächenfunktion sind (vgl. Abschnitt 3.6.1, Gleichung (3.37)), hat ein so konstruierter Spline die minimale Glattheit, die notwendig ist, um die Kraternormalen eindeutig zu interpolieren.

Zur Festlegung des freien Parameters  $\lambda$  wird das natürliche Maß der minimalen Biegeenergie der Oberfläche gewählt (formal ausgedrückt bedeutet das, dass das Integral über die Summe aller zweiten Ableitungen des resultierenden Splines  $s(\xi)$  minimal sein muss über alle  $\lambda$ ).

Das zur Konstruktion einer geeigneten Oberflächenfunktion  $\hat{r}$  zu lösende Problem lässt sich wie folgt formulieren:

„Finde denjenigen stückweise kubischen und überall mindestens einmal stetig differenzierbaren Spline, dessen Ableitung an den Stützstellen mit den zu den rekonstruierten Kraternormalen äquivalenten Gradienten übereinstimmt und die minimale Biegeenergie unter allen interpolierenden Splines aufweist.“

### 3.7.4 Triangulation

---

In den folgenden Abschnitten werden die mathematischen Grundlagendefinitionen gemacht, die nötig sind, um den Algorithmus zur Lösung dieses Problems zu formulieren. Der Schwerpunkt soll hier sein, die dem Algorithmus zugrundeliegenden Gleichungssysteme effizient aufzustellen. Es wird auf die Behandlung von Resultaten verzichtet, die nicht für den Anwendungsfall in dieser Arbeit relevant sind. Wo Zusammenhänge nicht sofort offensichtlich sind, wird auf die Literatur verwiesen werden.

Da man beim Vergleich der teilweise sehr alten und der neueren Literatur oft unvollständige, missverständliche oder widersprüchliche Indizierungsregeln und geometrische Konfigurationen der Dreiecke und Kontrollnetze der Splines antrifft, liegt der Schwerpunkt hier auf dem Bereitstellen von konsistenten Definitionen und korrekten Gleichungen. Ziel ist es, ein implementierbares Verfahren vorzustellen.

### 3.7.4 Triangulation

Der im Abschnitt 3.7.2 definierte Definitionsbereich der Oberflächenfunktion muss für die Konstruktion eines Splines geeignet unterteilt werden in Unterbereiche, die als Träger der Teilpolynome des Splines dienen. Die einfachstmögliche Art solcher Stücke sind die sogenannten *Simplizes*, und da  $\mathbb{D}$  per Konstruktion zweidimensional ist (Koordinaten  $\nu$  und  $\varphi$ ), sind die Simplizes hier Dreiecke.

Eine nicht-degenerierte<sup>6</sup> Punktmenge lässt sich immer derart mit Kanten überschneidungsfrei untereinander verbinden, dass eine Unterteilung der konvexen Hülle der Punktmenge in Dreiecke möglich ist, deren Vereinigung dann einfach zusammenhängend („lochfrei“) ist. Ein solcher spezieller Simplicialkomplex wird gemeinhin auch *Triangulation* einer Punktmenge genannt und lässt sich für den Anwendungsfall wie folgt definieren:

**Definition 3.7 (Triangulation der Krater)** Es seien  $\xi$  die  $N_p$  virtuellen Bildpunkte der erkannten Krater nach Definition 3.6 zu den Daten  $\mathbf{R}$  der räumlichen Rekonstruktion der Kraterparameter gemäß Def. 3.1a.

**a) Dreieck** Jedes Indextripel

$$t := (i, j, k) \text{ mit } 1 \leq i, j, k \leq N_p$$

---

<sup>6</sup> Alle Punkte müssen „in allgemeiner Lage“ liegen, dürfen also insbesondere nicht auf einer Geraden angeordnet sein.

wird Dreieck genannt, wenn die zu  $\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k \in \mathbf{R}$  korrespondierenden mittleren virtuellen Bildkoordinaten  $\xi_i, \xi_j, \xi_k \in \xi$  nicht kollinear liegen und zusätzlich positiv umlaufen. Das ist genau dann der Fall, wenn für die folgende Matrixdeterminante gilt:

$$\begin{vmatrix} \nu_i & \nu_j & \nu_k \\ \varphi_i & \varphi_j & \varphi_k \\ 1 & 1 & 1 \end{vmatrix} > 0.$$

**b) Triangulation** Eine Menge  $\Delta$  von Dreiecken  $t$  in diesem Sinne heißt Triangulation genau dann, wenn der Schnitt (bezogen auf die Punkte  $\xi$  und der auf ihnen basierenden Strecken und Flächen) je zweier verschiedener Dreiecke  $t$  und  $t'$  aus  $\Delta$  entweder leer oder eine gemeinsame Seite beider Dreiecke ist, und zusätzlich die Vereinigung aller Dreiecke aus  $\Delta$  genau die konvexe Hülle der Punkte  $\xi$  ist.

**c) Teiltriangulation** Eine Dreiecksmenge  $\Delta^{\subset}$  wird Teiltriangulation von  $\Delta$  genannt, wenn  $\Delta^{\subset} \subset \Delta$  ist und sie mindestens zusammenhängend ist.

**d) Triangulation als Index** Eine Triangulation  $\Delta$  bildet eindeutig auf einen geordneten Indexvektor  $\mathcal{I}_{\Delta}$  der Ecken-Indizes ihrer Dreiecke ab:

$$\mathcal{I}_{\Delta} := \text{idx} \left( \bigcup_{t \in \Delta} \{t_1, t_2, t_3\} \right),$$

wobei der Operator „idx“ seine Argumentmenge aufsteigend abzählt und das Tupel aller so geordneten Elemente zurückliefert.

Einige Bemerkungen zu dieser Definition: Die Triangulation  $\Delta$  ist ihre eigene maximale Teiltriangulation im Sinne von Def. 3.7c. Jedes einzelne Dreieck  $t \in \Delta$  ist eine minimale Teiltriangulation  $t =: \Delta^{\subset} \subset \Delta$ . Es existieren im Allgemeinen viele verschiedene Triangulationen der Punktmenge  $\xi$ . Für die folgende theoretische Konstruktion der Algorithmen ist die konkrete Wahl aber unerheblich. Bei Implementierung auf Rechnern benutzt man generell die sogenannte DELAUNAY-Triangulation, da diese numerische Vorteile und eine gewisse Eindeutigkeit bietet (in allen folgenden Beispielen und Diagrammen wird diese spezielle Triangulation ohne weiteren Kommentar verwendet werden).

In den Abbildungen 3.15 und 3.16 ist der Zusammenhang zwischen einem Beispielbild, darin erkannter Krater und der oben beschriebenen Triangulation ihrer Zentren visualisiert. Wie man durch Vergleich der Topologien auf der

### 3.7.5 BERNSTEIN-Polynome und BÉZIER-Splines

linken wie rechten Seite dort sieht, stellt die Konstruktion des virtuellen Bildkoordinatensystems sicher, dass die Triangulation nahezu verzerrungsfrei zwischen Bildkoordinaten und virtuellen Bildkoordinaten transformiert wird.

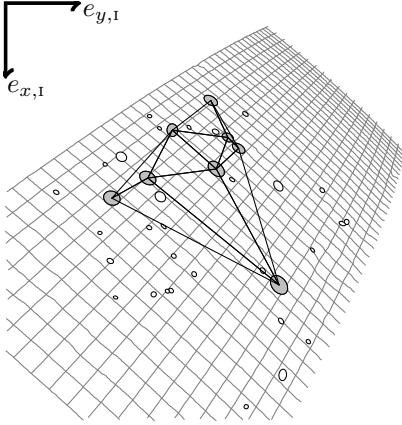


Abbildung 3.15: Die ins Bild transformierte Triangulation über die Zentren der größten erkannten Krater

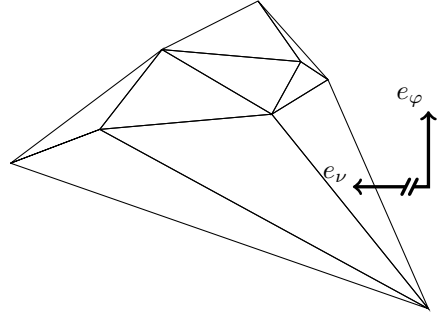


Abbildung 3.16: Transformation der Zentren in das virtuelle Bildkoordinatensystem und ihre Triangulation

### 3.7.5 BERNSTEIN-Polynome und BÉZIER-Splines

Die konvexe Hülle aller Kraterzentrenpunkte  $\xi \in \xi$  ist per der obigen Konstruktion „gepflastert“ mit den Dreiecken  $t \in \Delta$ . Die gesuchte Interpolante  $s(\xi)$  soll ausgedrückt werden durch hinreichend glatte Fortsetzung von auf den Dreiecken  $t$  niedergradigen Polynomen  $s^t$  auf die gesamte Triangulation.

Auf den jeweiligen Dreiecken  $t$  lassen sich Polynome  $n$ -ten Grades effizient in der sogenannten BERNSTEIN-BÉZIER-Form beschreiben, die als Linearkombination von Monomen der baryzentrischen Koordinaten auf den Dreiecken konstruiert werden.

**Definition 3.8 (BERNSTEIN-Polynome)** Sei  $\Delta$  eine Triangulation mit  $N_t$  Dreiecken der Zentren  $\xi = \{\xi_1, \dots, \xi_{N_p}\}$  der  $N_p$  erkannten Krater im Bild. Zu einem Dreieck  $t$  aus ihr seien definiert:

- a) **Baryzentrische Koordinaten** Ein Punkt  $\xi = (\nu, \varphi)^T \in \xi_t$  (vgl. Def.

3.6) besitzt in Bezug auf das Dreieck  $t$  die baryzentrischen Koordinaten

$$\boldsymbol{\lambda}^t(\xi) = (\lambda_1^t, \lambda_2^t, \lambda_3^t)^T$$

mit den Komponenten

$$\begin{aligned}\lambda_1^t(\xi) &= \frac{(\varphi_j - \varphi_k)(\nu_i - \nu_k) + (\nu_k - \nu_j)(\varphi_i - \varphi_k)}{(\varphi_j - \varphi_k)(\nu_i - \nu_k) + (\nu_k - \nu_j)(\varphi_i - \varphi_k)} \\ \lambda_2^t(\xi) &= \frac{(\varphi_k - \varphi_i)(\nu_i - \nu_k) + (\nu_i - \nu_k)(\varphi_i - \varphi_k)}{(\varphi_j - \varphi_k)(\nu_i - \nu_k) + (\nu_k - \nu_j)(\varphi_i - \varphi_k)} \\ \lambda_3^t(\xi) &= 1 - \lambda_1^t(\xi) - \lambda_2^t(\xi)\end{aligned}\quad (3.60)$$

für  $t = (i, j, k)$ . Der hochgestellte Index an  $\boldsymbol{\lambda}^t$  zeigt den Bezug auf  $t$  an.

**b) BERNSTEIN-Polynom** Ein Polynom  $s^t(\boldsymbol{\lambda}^t)$  vom Grad 3 in der Fläche des Dreiecks  $t$  von der Form

$$s^t(\boldsymbol{\lambda}^t) = \begin{cases} \sum_{|\alpha|=3} c_\alpha^t B_\alpha^t(\boldsymbol{\lambda}^t), & \text{wenn } 0 \leq \lambda_1, \lambda_2, \lambda_3 \leq 1 \\ 0, & \text{sonst} \end{cases}\quad (3.61)$$

mit  $\alpha \in \mathbb{N}^3$  und

$$B_\alpha^t(\boldsymbol{\lambda}^t) = \frac{|\alpha|!}{\alpha!} \boldsymbol{\lambda}^{t\alpha}\quad (3.62)$$

heißt BERNSTEIN-Polynom.  $c_\alpha^t$  sind die Koeffizienten des Polynoms,  $\alpha$  ein Multiindex mit den üblichen Konventionen bezüglich Fakultät, Betrag, Exponentiation und Indizierung.

Nach der Definition 3.8b ist der Träger von  $s^t$  also genau das Dreieck  $t$ , da außerhalb mindestens eine der Komponenten von  $\boldsymbol{\lambda}$  wegen der Definition 3.8a größer eins oder kleiner null wäre. Die Koordinaten  $\boldsymbol{\lambda}^t$  und die über sie gebildeten Polynome  $B^t(\boldsymbol{\lambda}^t)$  beziehen<sup>7</sup> sich immer auf die Ecken des Dreiecks  $t$ .

Die baryzentrischen Koordinaten für ein Dreieck über drei beliebige Punkte  $\xi_i$ ,  $\xi_j$  und  $\xi_k$  der virtuellen Bildkoordinaten sind illustriert in Abbildung 3.17. Dort sind auch die Drittel-Isolinien eingezeichnet, deren Schnittpunkte miteinander und mit den Dreieckseiten die Positionen der sogenannten „Kontrollpunkte“ anzeigen.

Es besteht ein direkter Zusammenhang zwischen den baryzentrischen

<sup>7</sup>Wo mehrere Symbole mit dem  $t$ -Index in einer Gleichung, wie z.B. in Gleichung (3.61) der Definition 3.8, vorkommen, können bei Unmissverständlichkeit redundante Indizes  $t$  der besseren Lesbarkeit halber weggelassen werden.

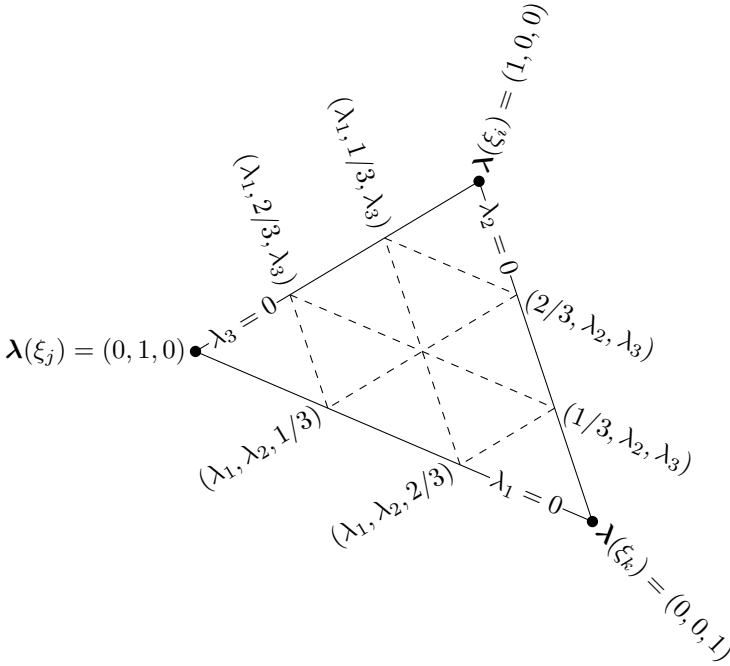


Abbildung 3.17: Baryzentrische Koordinaten mit Isolinen auf einem Dreieck

Koordinaten der Schnittpunkte der Isolinen in Abbildung 3.17 und den Multiindizes  $\alpha$  der Koeffizienten der BERNSTEIN-Polynome. Dieser ist in der Tabelle 3.1 iteriert. Die korrespondierenden baryzentrischen Koordinaten der Indizes  $\alpha_1, \dots, \alpha_{10}$  aus der Tabelle 3.1 und die durch die  $\alpha$  indizierten Polynomkoeffizienten  $c_{\alpha}^t$  bilden das sogenannte „Kontrollnetz“ der BERNSTEIN-Polynome. Dieses ist am Beispiel des selben Dreiecks in Abbildung 3.18 gezeigt. Man erkennt die Kongruenz mit den Drittel-Isolinen der baryzentrischen Koordinaten in Abbildung 3.17. Der Name „Kontrollpunkt“ rührt daher, dass an diesen Punkten der Wert des BERNSTEIN-Polynoms mit dem Wert seines dortigen Koeffizienten übereinstimmt:

$$s^t\left(\frac{\alpha_i}{3}; c^t\right) = c_{\alpha_i}^t. \quad (3.63)$$

(Wie sich leicht anhand der Definition 3.8b überprüfen lässt.) Die gesuch-

$\alpha$ -#	$\alpha$	$\lambda$
1	(0,0,3)	(0, 0, 1)
2	(0,1,2)	$(0, \frac{1}{3}, \frac{2}{3})$
3	(0,2,1)	$(0, \frac{2}{3}, \frac{1}{3})$
4	(0,3,0)	(0, 1, 0)
5	(1,0,2)	$(\frac{1}{3}, 0, \frac{2}{3})$
6	(1,1,1)	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$
7	(1,2,0)	$(\frac{1}{3}, \frac{2}{3}, 0)$
8	(2,0,1)	$(\frac{2}{3}, 0, \frac{1}{3})$
9	(2,1,0)	$(\frac{2}{3}, \frac{1}{3}, 0)$
10	(3,0,0)	(1, 0, 0)
$\Sigma = 3$		$\Sigma = 1$

Tabelle 3.1: Zusammenhang der Multiindizes mit den baryzentrischen Kontrollpunkten

te Spline-Funktion  $s$  auf der Triangulation  $\Delta$  wird aus diesen einzelnen Polynomen auf den Teildreiecken konstruiert und wird als BÉZIER-Spline bezeichnet:

**Definition 3.9 (BÉZIER-Spline)** Sei  $\xi = \{\xi_1, \dots, \xi_{N_p}\} \subset \mathbb{D}$  die Menge der mittleren Punkte der erkannten Kraterzentren und  $\Delta$  eine Triangulation über  $\xi$ . Die Funktion

$$s(\xi) := \sum_{t \in \Delta} s^t(\lambda(\xi)), \quad \xi \in \mathbb{D} \quad (3.64)$$

wird der aus den BERNSTEIN-Polynomen dritten Grades  $s^t$  konstruierte kubische, einmal stetig differenzierbare BÉZIER-Spline genannt, wenn für je zwei Dreiecke  $t, t' \in \Delta$ , die eine gemeinsame Seite besitzen, gilt

$$\begin{aligned} s^t(\lambda) &= s^{t'}(\lambda) & (\text{Stetigkeit}) \\ D^\beta s^t(\lambda) &= D^\beta s^{t'}(\lambda) & (\text{stetige Differenzierbarkeit}) \end{aligned} \quad (3.65)$$

für alle  $\lambda$  auf dem Schnitt von  $t$  und  $t'$ .  $\beta \in \mathbb{N}^3$  ist Multiindex mit  $|\beta| = 1$  und  $D^\beta$  die Ableitung nach  $\lambda$ .

Die Übergangsbedingungen (3.65) sind zwar kontinuierlich, Farin zeigte aber

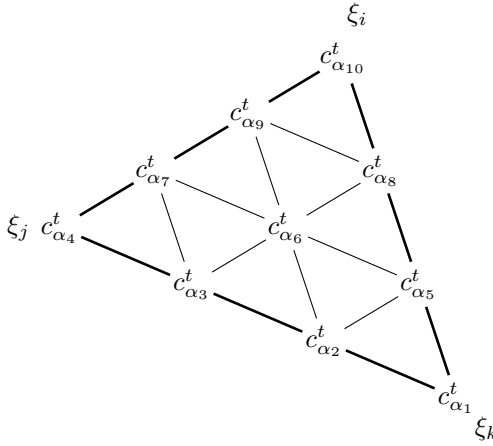



Abbildung 3.18: Das Kontrollnetz des BERNSTEIN-Polynoms vom Grad 3

in seiner Dissertation, dass sie äquivalent zu einfachen linearen Zusammenhängen der Koeffizienten  $c_{\alpha}^t$  sind (Farin, 1979). So stellte sich heraus, dass die Stetigkeitsbedingung bereits erfüllt ist, wenn die gemeinsamen Koeffizienten der beiden Polynome identisch sind, wie in Abbildung 3.19 illustriert<sup>8</sup> für zwei Dreiecke  $t$  über  $i, j, k$  und  $t'$  über  $i, k, l$ .

Auch für die Bedingung der stetigen Differenzierbarkeit gibt es einen einfachen geometrischen Zusammenhang: Die an der  $t$  und  $t'$  gemeinsamen Seite aneinanderstoßenden durch die Kontrollpunkte gebildeten Unterdreiecke müssen koplanar sein, um die stetige Differenzierbarkeit zu gewährleisten (Farin, 1986). Die beteiligten Kontrollpunkte sind am Beispiel des Übergangs von  $t'$  nach  $t$  in Abbildung 3.20 und der Tabelle 3.2 illustriert. Die Bedingungen sind natürlich symmetrisch und könnten auch für den Übergang  $t$  nach  $t'$  aufgestellt werden.

Das Zahnrad  steht dort für die Beziehungen, die in den Tabellen 3.2 und 3.3 aufgelistet sind, wobei  $\xi_t = (\xi_i, \xi_j, \xi_k)$  in dem Zusammenhang das linke Dreieck („L“) ist und mit Seite 2 (zweite baryzentrische Koordinate, da gegenüber der zweiten Ecke  $\xi_j$ ) an das rechte Dreieck („R“)  $\xi_{t'} = (\xi_i, \xi_k, \xi_l)$  an dessen Seite 3 (dritte baryzentrische Koordinate zu  $\xi_l$ ) stößt. Es handelt sich also um Fall „L2-R3“. Die Größe  $\lambda^R(\xi^L)$  in den Übergangsgleichungen

---

<sup>8</sup>Der Lesbarkeit halber werden statt der Bezeichner „ $c_{\alpha_i}^t$ “ die korrespondierenden Indizes der ersten Spalte aus Tabelle 3.1 verwendet.



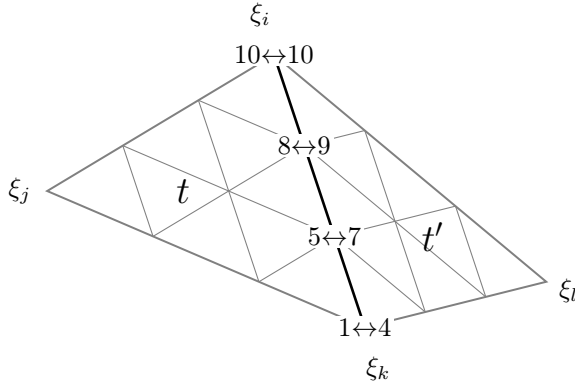


Abbildung 3.19: Geometrische Stetigkeitsbedingungen eines Splines

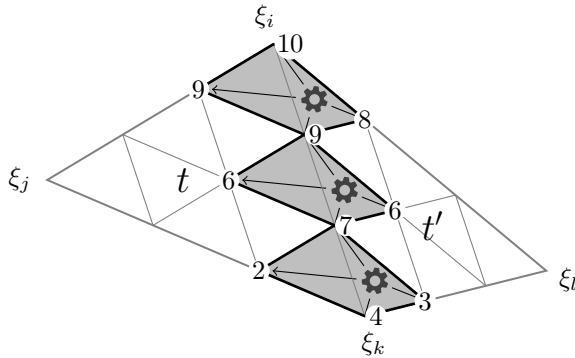


Abbildung 3.20: Geometrische Bedingungen zur stetigen Differenzierbarkeit eines Splines

beschreibt die baryzentrischen Koordinaten der nicht-gemeinsamen Ecke  $\xi_j$  des linken Dreiecks bezüglich den Ecken  $\xi_i$ ,  $\xi_k$ ,  $\xi_l$  des rechten Dreiecks (es wird also  $\xi_j$  als freier Punkt in die Gleichungen (3.60) eingesetzt).

Bezeichnet man die drei Komponenten von  $\lambda^R(\xi^L)$  kurz als  $\lambda_1$ ,  $\lambda_2$  und  $\lambda_3$ , dann lassen sich die Gleichungssysteme in der Tabelle 3.3 als in den Kontrollpunktvektoren  $c^L$  und  $c^R$  lineare Gleichungen schreiben. Durch Definition der

Matrizen

$$\begin{aligned}
C^{L3} &:= \begin{pmatrix} 0010000000 \\ 0000010000 \\ 0000000100 \end{pmatrix}, \quad C^{L2} := \begin{pmatrix} 0000000010 \\ 0000010000 \\ 0100000000 \end{pmatrix}, \\
C^{L1} &:= \begin{pmatrix} 0000100000 \\ 0000010000 \\ 0000001000 \end{pmatrix},
\end{aligned} \tag{3.66}$$

und

$$\begin{aligned}
C^{R3} &:= \begin{pmatrix} 00\ 0\ 0\ 0\ 0\ 0\ \lambda_3\ \lambda_2\ \lambda_1 \\ 00\ 0\ 0\ 0\ \lambda_3\ \lambda_2\ 0\ \lambda_1\ 0 \\ 00\ \lambda_3\ \lambda_2\ 0\ 0\ 0\ \lambda_1\ 0\ 0 \end{pmatrix}, \quad C^{R2} := \begin{pmatrix} \lambda_3\ \lambda_2\ 00\ \lambda_1\ 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 00\ \lambda_3\ \lambda_2\ 0\ \lambda_1\ 0\ 0 \\ 0\ 0\ 00\ 0\ 0\ 0\ \lambda_3\ \lambda_2\ \lambda_1 \end{pmatrix}, \\
C^{R1} &:= \begin{pmatrix} 0\ 0\ \lambda_3\ \lambda_2\ 0\ 0\ \lambda_1\ 000 \\ 0\ \lambda_3\ \lambda_2\ 0\ 0\ \lambda_1\ 0\ 000 \\ \lambda_3\ \lambda_2\ 0\ 0\ \lambda_1\ 0\ 0\ 000 \end{pmatrix},
\end{aligned} \tag{3.67}$$

können alle Gleichungen der Tabelle 3.3 ausgedrückt werden als

$$\begin{aligned}
\text{L3-R3} &\Leftrightarrow C^{L3}c^L = C^{R3}c^R \\
\text{L3-R2} &\Leftrightarrow C^{L3}c^L = C^{R2}c^R \\
&\vdots
\end{aligned} \tag{3.68}$$

Für je eine geteilte Seite zweier Dreiecke liegen damit sieben Übergangsbedingungen vor: Vier aus den Stetigkeitsbedingungen ( $\mathcal{C}^0$ ) und drei aus den Bedingungen zu stetiger Differenzierbarkeit ( $\mathcal{C}^1$ ). Aufgrund der Symmetrie sind die Bedingungen bei Vertauschung der Dreiecke redundant und müssen nicht zusätzlich in die später zu lösenden Gleichungssysteme aufgenommen werden. Für die Kontrollpunkte am Rand der Triangulation gibt es keine Bedingungen. Man sagt, der Spline „besitzt eine freie Randbedingung“.

Fall	$\mathcal{C}^0$	$\mathcal{C}^1$
L3-R3	$c_{(i,j,0)}^L = c_{(j,i,0)}^R$	$c_{(i,j,1)}^L = \sum_{ \gamma =1} c_{(j,i,0)+\gamma}^R B_\gamma(\lambda^R(\xi^L))$
L3-R2	$c_{(i,j,0)}^L = c_{(i,0,j)}^R$	$c_{(i,j,1)}^L = \sum_{ \gamma =1} c_{(i,0,j)+\gamma}^R B_\gamma(\lambda^R(\xi^L))$
L3-R1	$c_{(i,j,0)}^L = c_{(0,j,i)}^R$	$c_{(i,j,1)}^L = \sum_{ \gamma =1} c_{(0,j,i)+\gamma}^R B_\gamma(\lambda^R(\xi^L))$
L2-R3	$c_{(i,0,j)}^L = c_{(i,j,0)}^R$	$c_{(i,1,j)}^L = \sum_{ \gamma =1} c_{(i,j,0)+\gamma}^R B_\gamma(\lambda^R(\xi^L))$
L2-R2	$c_{(i,0,j)}^L = c_{(j,0,i)}^R$	$c_{(i,1,j)}^L = \sum_{ \gamma =1} c_{(j,0,i)+\gamma}^R B_\gamma(\lambda^R(\xi^L))$
L2-R1	$c_{(i,0,j)}^L = c_{(0,i,j)}^R$	$c_{(i,1,j)}^L = \sum_{ \gamma =1} c_{(0,i,j)+\gamma}^R B_\gamma(\lambda^R(\xi^L))$
L1-R3	$c_{(0,i,j)}^L = c_{(j,i,0)}^R$	$c_{(1,i,j)}^L = \sum_{ \gamma =1} c_{(j,i,0)+\gamma}^R B_\gamma(\lambda^R(\xi^L))$
L1-R2	$c_{(0,i,j)}^L = c_{(i,0,j)}^R$	$c_{(1,i,j)}^L = \sum_{ \gamma =1} c_{(i,0,j)+\gamma}^R B_\gamma(\lambda^R(\xi^L))$
L1-R1	$c_{(0,i,j)}^L = c_{(0,j,i)}^R$	$c_{(1,i,j)}^L = \sum_{ \gamma =1} c_{(0,j,i)+\gamma}^R B_\gamma(\lambda^R(\xi^L))$

Tabelle 3.2:  $\mathcal{C}^0$ - und  $\mathcal{C}^1$ -Übergangsbedingungen in Abhängigkeit von den geteilten Seiten der Dreiecke von rechts nach links.

### 3.7.5 BERNSTEIN-Polynome und BÉZIER-Splines

Fall	Übergangsgleichungen				$\alpha$ -Anordnung
L3-R3	$c_3^L = c_{10}^R \lambda_1^R + c_9^R \lambda_2^R + c_8^R \lambda_3^R$	+			
	$c_6^L = c_9^R \lambda_1^R + c_7^R \lambda_2^R + c_6^R \lambda_3^R$	+			
	$c_8^L = c_7^R \lambda_1^R + c_4^R \lambda_2^R + c_3^R \lambda_3^R$	+			
L3-R2	$c_3^L = c_5^R \lambda_1^R + c_2^R \lambda_2^R + c_1^R \lambda_3^R$	+			
	$c_6^L = c_8^R \lambda_1^R + c_6^R \lambda_2^R + c_5^R \lambda_3^R$	+			
	$c_8^L = c_{10}^R \lambda_1^R + c_9^R \lambda_2^R + c_8^R \lambda_3^R$	+			
L3-R1	$c_3^L = c_7^R \lambda_1^R + c_4^R \lambda_2^R + c_3^R \lambda_3^R$	+			
	$c_6^L = c_6^R \lambda_1^R + c_3^R \lambda_2^R + c_2^R \lambda_3^R$	+			
	$c_8^L = c_5^R \lambda_1^R + c_2^R \lambda_2^R + c_1^R \lambda_3^R$	+			
L2-R3	$c_9^L = c_{10}^R \lambda_1^R + c_9^R \lambda_2^R + c_8^R \lambda_3^R$	+			
	$c_6^L = c_9^R \lambda_1^R + c_7^R \lambda_2^R + c_6^R \lambda_3^R$	+			
	$c_2^L = c_7^R \lambda_1^R + c_4^R \lambda_2^R + c_3^R \lambda_3^R$	+			
L2-R2	$c_9^L = c_5^R \lambda_1^R + c_2^R \lambda_2^R + c_1^R \lambda_3^R$	+			
	$c_6^L = c_8^R \lambda_1^R + c_6^R \lambda_2^R + c_5^R \lambda_3^R$	+			
	$c_2^L = c_{10}^R \lambda_1^R + c_9^R \lambda_2^R + c_8^R \lambda_3^R$	+			
L2-R1	$c_9^L = c_7^R \lambda_1^R + c_4^R \lambda_2^R + c_3^R \lambda_3^R$	+			
	$c_6^L = c_6^R \lambda_1^R + c_3^R \lambda_2^R + c_2^R \lambda_3^R$	+			
	$c_2^L = c_5^R \lambda_1^R + c_2^R \lambda_2^R + c_1^R \lambda_3^R$	+			
L1-R3	$c_5^L = c_{10}^R \lambda_1^R + c_9^R \lambda_2^R + c_8^R \lambda_3^R$	+			
	$c_6^L = c_9^R \lambda_1^R + c_7^R \lambda_2^R + c_6^R \lambda_3^R$	+			
	$c_7^L = c_7^R \lambda_1^R + c_4^R \lambda_2^R + c_3^R \lambda_3^R$	+			
L1-R2	$c_5^L = c_5^R \lambda_1^R + c_2^R \lambda_2^R + c_1^R \lambda_3^R$	+			
	$c_6^L = c_8^R \lambda_1^R + c_6^R \lambda_2^R + c_5^R \lambda_3^R$	+			
	$c_7^L = c_{10}^R \lambda_1^R + c_9^R \lambda_2^R + c_8^R \lambda_3^R$	+			
L1-R1	$c_5^L = c_7^R \lambda_1^R + c_4^R \lambda_2^R + c_3^R \lambda_3^R$	+			
	$c_6^L = c_6^R \lambda_1^R + c_3^R \lambda_2^R + c_2^R \lambda_3^R$	+			
	$c_7^L = c_5^R \lambda_1^R + c_2^R \lambda_2^R + c_1^R \lambda_3^R$	+			

Tabelle 3.3: Explizite  $\mathcal{C}^1$ -Übergangsbedingungen mit Skizzen der Anordnung der Kontrollnetz-Indizes  $\alpha$ -#

### 3.7.6 Interpolationseigenschaft

Der im vorangegangenen Abschnitt konstruierte Spline mit seinen inneren Übergangsbedingungen muss schließlich an seinen Stützstellen (den erkannten Kraterzentren  $\xi \in \mathfrak{X}$ ) eine Interpolationsbedingung erfüllen. Den Argumenten des Abschnitts 3.6.2 folgend muss an diesen Stellen die Ableitung des Splines in  $\nu$ - und  $\varphi$ -Richtung mit den zu den rekonstruierten Kraternormalen äquivalenten  $\nu$ - $\varphi$ -Gradienten übereinstimmen.

Stellt man also einen Spline auf, dessen Ableitung an den Stellen  $\xi_i = (\nu_i, \varphi_i)^T$  die skalierten Gradienten aus Gleichung (3.53) interpoliert, so beschreibt dieser resultierende Spline den Integralterm in der Exponentialfunktion auf der rechten Seite der Gleichung (3.54). Das Exponential jedes solchen Splines ist dann per Konstruktion eine zur Definition 3.2 konforme Oberfläche  $\hat{r}(\nu, \varphi)$ .

Der skalierte Gradient  $\nabla r(\xi)/r(\xi)$  in virtuellen Bildkoordinaten ergibt sich aus den Daten  $\mathbf{R}$  durch den Zusammenhang (3.53). Diese Gleichung enthält die konfigurierten Zentrums- und Normalenvektoren  $(\hat{c}_i, \hat{n}_i) = R_{i, \kappa_i}$ , sie ist also implizit von der Konfiguration  $\kappa$  abhängig.

Um die Gradienten-Interpolationsgleichungen zu formulieren, muss auf einem Dreieck  $t$ , an dessen Ecke eine Interpolationsbedingung vorliegt, das baryzentrische BERNSTEIN-Polynom  $s^t$  (vgl. Definition 3.8b) ebenfalls in den virtuellen Bildkoordinaten  $\xi = (\nu, \varphi)^T$  ausgedrückt werden.

Definiert man für die vorkommenden Differenzen in Gleichung (3.60) der Definition 3.8a

$$\begin{aligned} a &:= \varphi_2 - \varphi_3, & b &:= \nu_3 - \nu_2, \\ c &:= \varphi_3 - \varphi_1, & d &:= \nu_1 - \nu_3, \\ e &:= ad - bc, \end{aligned} \tag{3.69}$$

(1, 2, 3 sind die Indizes der Ecken von  $t$ ), und damit

$$\mathbf{\Gamma}^t := \frac{1}{e} \begin{pmatrix} a & b \\ c & d \\ -(a+c) & -(b+d) \end{pmatrix}, \tag{3.70}$$

wird die Transformation von virtuellen Bild- nach baryzentrischen Koordinaten zur Matrixoperation

$$\lambda^t(\xi) = \mathbf{\Gamma}^t \begin{pmatrix} \nu - \nu_3 \\ \varphi - \varphi_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.71}$$

### 3.7.6 Interpolationseigenschaft

---

Da die Triangulation über die *mittleren virtuellen Kraterkoordinaten* (vgl. Def. 3.6) gebildet wurde, hängt die Koeffizientenmatrix  $\mathbf{\Gamma}$  gerade nicht von der Konfiguration  $\kappa$  ab. Den dadurch entstehenden Approximationsfehler abzuschätzen ist Gegenstand eines späteren Kapitels.

Die JACOBI-Matrix von  $\lambda^t$  bezüglich  $\xi$  ist wegen der Linearität von (3.71) wieder die konstante Matrix

$$J_\xi[\lambda^t] = \mathbf{\Gamma}^t. \quad (3.72)$$

Die JACOBI-Matrix des BERNSTEIN-Polynoms  $s^t$  bezüglich der baryzentrischen Koordinaten auf  $t$  ist wegen (3.61)

$$J_\lambda[s^t](\lambda) = \sum_{|\alpha|=3} c_\alpha^t J_\lambda[B_\alpha^t](\lambda). \quad (3.73)$$

Die beiden Gleichungen (3.72) und (3.73) ergeben dann zusammen mit der Kettenregel die Ableitung des BERNSTEIN-Polynoms  $s^t$  nach den virtuellen Bildkoordinaten  $\xi$ :

$$\begin{aligned} J_\xi[s^t \circ \lambda](\xi) &\stackrel{\text{Kettenr.}}{=} J_\lambda[s^t](\lambda(\xi)) \cdot J_\xi[\lambda](\xi) \\ &\stackrel{(3.72)}{=} J_\lambda[s^t](\lambda(\xi)) \cdot \mathbf{\Gamma}^t \\ &\stackrel{(3.73)}{=} \sum_{|\alpha|=3} c_\alpha^t J_\lambda[B_\alpha^t](\lambda(\xi)) \cdot \mathbf{\Gamma}^t \\ &\stackrel{\text{Kettenr.}}{=} \sum_{|\alpha|=3} c_\alpha^t J_\xi[B_\alpha^t \circ \lambda](\xi) \end{aligned} \quad (3.74)$$

Definiere

$$\gamma_1 := (1, 0, 0), \quad \gamma_2 := (0, 1, 0), \quad \gamma_3 := (0, 0, 1), \quad (3.75)$$

dann ist die  $\lambda$ -JACOBI-Matrix des Faktors  $B_\alpha^t$  in den obenstehenden Gleichungen

$$\begin{aligned} J_\lambda[B_\alpha^t](\lambda) &= \frac{|\alpha|!}{\alpha!} (D_\lambda^{\gamma_1} B_\alpha^t(\lambda), D_\lambda^{\gamma_2} B_\alpha^t(\lambda), D_\lambda^{\gamma_3} B_\alpha^t(\lambda)) \\ &= \frac{|\alpha|!}{\alpha!} (\alpha_1 \lambda^{\alpha-\gamma_1}, \alpha_2 \lambda^{\alpha-\gamma_2}, \alpha_3 \lambda^{\alpha-\gamma_3}). \end{aligned} \quad (3.76)$$

Die Indizes an  $\alpha$  zeigen hier die Komponenten des Multiindex an, nicht den linearen Index aus der Tabelle 3.1.

Man sieht, dass die Ableitungen nach den baryzentrischen Koordinaten einen Vektor wiederum niedergradiger BERNSTEIN-Polynome beschreibt<sup>9</sup>. Alle  $\xi$ -JACOBI- $(1 \times 2)$ -Matrizen der  $B_\alpha^t$  in der letzten Zeile der Gleichung (3.74) sind untereinander in Tabelle 3.4 aufgelistet<sup>10</sup>. Mit ihnen sei die folgende  $(2 \times 10)$ -„Ableitungsmatrix“

$$\begin{aligned} B_\xi^t(\xi) &:= \left( (J_\xi[B_{\alpha_1}^t \circ \lambda](\xi))^T, \dots, (J_\xi[B_{\alpha_{10}}^t \circ \lambda](\xi))^T \right) \\ &= \begin{pmatrix} D_\xi^{(1,0)} B_{\alpha_1}^t(\xi) & \dots & D_\xi^{(1,0)} B_{\alpha_{10}}^t(\xi) \\ D_\xi^{(0,1)} B_{\alpha_1}^t(\xi) & \dots & D_\xi^{(0,1)} B_{\alpha_{10}}^t(\xi) \end{pmatrix} \end{aligned} \quad (3.77)$$

definiert, sowie die zu den Kraternormalen äquivalenten  $(2 \times 1)$ -Gradienten in den virtuellen Bildkoordinaten nach Gleichung (3.53)

$$G_i^\kappa := \left( \frac{\nabla r(\nu_i, \varphi_i)}{r(\nu_i, \varphi_i)} \right)^T \quad (3.78)$$

und der  $(10 \times 1)$ -Koeffizientenvektor

$$c^t := (c_{\alpha_1}^t, \dots, c_{\alpha_{10}}^t)^T. \quad (3.79)$$

Der Konfigurationsparameter  $\kappa$  in  $G_i^\kappa$  in (3.78) kann im Folgenden weggelassen werden, da er für das Prinzip der Konstruktion unerheblich ist. Er wird erst in späteren Abschnitten interessant, wenn der Vergleich verschiedener Konfigurationen Gegenstand ist. In den Gleichungen (3.77) und (3.79) vorkommende Indizes an  $\alpha$  beziehen sich auf deren lineare Nummerierung und nicht auf ihre Komponenten. Die Matrix  $B_\xi^t$  ist praktisch die transponierte Tabelle 3.4.

Zu jeder Ecke  $\xi_i$  jedes Dreieckes  $t \in \Delta$  gibt es mit den obigen drei Definitionen dann zwei lineare Interpolationsbedingungen:

$$B_\xi^t(\xi_i) \cdot c^t = G_i \quad (3.80)$$

Vorteilhaft ist, dass die Gleichungen (3.80) per Definition nur an den Ecken der Dreiecke ausgewertet werden müssen. Daher sind für alle  $\xi_i = (\nu_i, \varphi_i)^T$

<sup>9</sup>Dieser Umstand trifft tatsächlich für alle Ableitungen solcher Polynome zu (Farin, 1986)

<sup>10</sup>Die  $D_\xi^{(1,0)} B_\alpha^t$  und  $D_\xi^{(0,1)} B_\alpha^t$  bezeichnen jeweils die Ableitungen nach der ersten Komponente von  $\xi$ , also  $\nu$ , und der zweiten, also  $\varphi$ . Die Indizes  $t$  sind an den Koeffizienten  $a$ ,  $b$ ,  $c$ ,  $d$  und  $e$  der Lesbarkeit halber weggelassen, bei Gleichungssystemen über mehrere Dreiecke müssen sie berücksichtigt werden (vgl. Abbildung A.10 in Beispiel A.3)

### 3.7.7 Biegeenergie-Funktion

---

die baryzentrischen Koordinaten  $\lambda(\xi_i)$  entweder  $(1, 0, 0)$ ,  $(0, 1, 0)$  oder  $(0, 0, 1)$  und die Matrix  $\mathbf{B}_\xi^t$  nimmt nur die folgenden drei Werte an:

$$\mathbf{B}_\xi^t(1, 0, 0) = \frac{3}{e_t} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & -(a_t + c_t) & c_t & a_t \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -(b_t + d_t) & d_t & b_t \end{pmatrix}, \quad (3.81)$$

$$\mathbf{B}_\xi^t(0, 1, 0) = \frac{3}{e_t} \begin{pmatrix} 0 & 0 & -(a_t + c_t) & c_t & 0 & 0 & a_t & 0 & 0 & 0 \\ 0 & 0 & -(b_t + d_t) & d_t & 0 & 0 & b_t & 0 & 0 & 0 \end{pmatrix}, \quad (3.82)$$

und

$$\mathbf{B}_\xi^t(0, 0, 1) = \frac{3}{e_t} \begin{pmatrix} -(a_t + c_t) & c_t & 0 & 0 & a_t & 0 & 0 & 0 & 0 & 0 \\ -(b_t + d_t) & d_t & 0 & 0 & b_t & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (3.83)$$

(Man überprüfe durch Einsetzen in die Tabelle 3.4 auf Seite 98.)

### 3.7.7 Biegeenergie-Funktion

Zur Lösung des in Abschnitt 3.7.3 formulierten Minimierungsproblems soll jetzt dessen Zielfunktion, die Biegeenergie-Funktion des Splines, konstruiert werden. Als „Biegeenergie“ einer Funktion wird generell das Integral über die Summe des Quadrats ihrer zweiten partiellen Ableitungen bezeichnet, gebildet über den Bereich, für den diese Energie bestimmt werden soll. Bezeichnet man den Ableitungs-Multiindex als

$$\beta \in \{(2, 0), (1, 1), (1, 1), (0, 2)\}, \quad (3.84)$$

also  $\beta \in \mathbb{N}^2$  und  $|\beta| = 2$ , so ist die gesuchte Zielfunktion

$$E(s) = \sum_{t \in \Delta} \int_t \sum_{|\beta|=2} \left( D_\xi^\beta s^t \right)^2 \quad (3.85)$$

Die partiellen zweiten Ableitungen der BERNSTEIN-Polynome  $s^t$  im Integranden lassen sich aus deren HESSE-Matrix ablesen, die sich ähnlich wie (3.74)



aus der Kettenregel ableitet:

$$\begin{aligned}
 H_\xi[s^t \circ \lambda](\xi) &= J_\xi[(J_\xi[s^t \circ \lambda](\xi))^T](\xi) \\
 &= J_\xi[(J_\lambda[s^t](\lambda(\xi)) \cdot \Gamma^t)^T](\xi) \\
 &= J_\xi[(\Gamma^t)^T \cdot (J_\lambda[s^t](\lambda(\xi)))^T](\xi) \\
 &= (\Gamma^t)^T \cdot J_\xi[(J_\lambda[s^t](\lambda(\xi)))^T](\xi) \\
 &= (\Gamma^t)^T \cdot J_\xi[(J_\lambda[s^t])^T \circ \lambda](\xi) \\
 &= (\Gamma^t)^T \cdot J_\lambda[(J_\lambda[s^t])^T](\lambda(\xi)) \cdot \Gamma^t \\
 &= (\Gamma^t)^T \cdot H_\lambda[s^t](\lambda(\xi)) \cdot \Gamma^t \\
 &= \underbrace{(\Gamma^t)^T}_{2 \times 3} \cdot \underbrace{\left( \sum_{|\alpha|=3} c_\alpha^t H_\lambda[B_\alpha^t](\lambda(\xi)) \right)}_{\substack{3 \times 3 \\ 2 \times 2}} \cdot \underbrace{\Gamma^t}_{3 \times 2}
 \end{aligned} \tag{3.86}$$

Mit den  $\gamma_{1,2,3}$  wie definiert in Gleichung (3.75) lassen sich die HESSE-Matrizen der Faktoren  $B_\alpha^t$  aus der letzten Zeile der obigen Gleichung wie die JACOBI-Matrizen in Gleichung (3.76) als BERNSTEIN-Polynome (jetzt vom Grad 1) schreiben:

$$H_\lambda[B_\alpha^t](\lambda) = \begin{pmatrix} \alpha_1(\alpha_1 - 1)\lambda^{\alpha-2\gamma_1} & \alpha_1\alpha_2\lambda^{\alpha-\gamma_1-\gamma_2} & \alpha_1\alpha_3\lambda^{\alpha-\gamma_1-\gamma_3} \\ \alpha_1\alpha_2\lambda^{\alpha-\gamma_1-\gamma_2} & \alpha_2(\alpha_2 - 1)\lambda^{\alpha-2\gamma_2} & \alpha_2\alpha_3\lambda^{\alpha-\gamma_2-\gamma_3} \\ \alpha_1\alpha_3\lambda^{\alpha-\gamma_1-\gamma_3} & \alpha_2\alpha_3\lambda^{\alpha-\gamma_2-\gamma_3} & \alpha_3(\alpha_3 - 1)\lambda^{\alpha-2\gamma_3} \end{pmatrix} \tag{3.87}$$

Die letzte Zeile von Gleichung(3.86), die Summe der Matrizen aus (3.87) für alle  $\alpha$ , links-rechts-multipliziert mit  $\Gamma^t$ , bildet also die  $2 \times 2$ - $\xi$ -HESSE-Matrix von  $s^t$ . Die drei nichtsymmetrischen Elemente ihrer einzelnen Summanden können in Tabelle 3.5 auf Seite 99 nachgeschlagen werden. Sei ähnlich wie in (3.77) aus diesen Summanden die Matrix der zweiten Ableitungen

$$\mathbf{B}_{\xi\xi}^t := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} D_\xi^{(2,0)} B_{\alpha_1}^t(\xi) & D_\xi^{(2,0)} B_{\alpha_{10}}^t(\xi) \\ D_\xi^{(1,1)} B_{\alpha_1}^t(\xi) & \cdots & D_\xi^{(1,1)} B_{\alpha_{10}}^t(\xi) \\ D_\xi^{(0,2)} B_{\alpha_1}^t(\xi) & D_\xi^{(0,2)} B_{\alpha_{10}}^t(\xi) \end{pmatrix} \tag{3.88}$$

definiert und die  $B_\alpha$  selbst, wie ihre Ableitungen in obiger Gleichung, in einem Vektor  $\mathbf{B}^t := (B_{\alpha_1}^t, \dots, B_{\alpha_{10}}^t)^T$  zusammengefasst. Dann vereinfachen sich

### 3.7.7 Biegeenergie-Funktion

zusammen mit dem Koeffizientenvektor  $c^t$  aus Gleichung (3.79) die Integrale über die Dreiecke  $t$  in Gleichung (3.85) wie folgt:

$$\begin{aligned}
 \int_t \sum_{|\beta|=2} \left( D_\xi^\beta s^t \right)^2 &= \int_t \sum_{|\beta|=2} \left( D_\xi^\beta \sum_{|\alpha|=3} c_\alpha^t B_\alpha^t \right)^2 \\
 &= \int_t \sum_{|\beta|=2} \left( \sum_{|\alpha|=3} c_\alpha^t D_\xi^\beta B_\alpha^t \right)^2 \\
 &= \int_t \left( \begin{aligned} &\left( \sum_{|\alpha|=3} c_\alpha^t D_\xi^{(2,0)} B_\alpha^t \right)^2 \\ &+ 2 \left( \sum_{|\alpha|=3} c_\alpha^t D_\xi^{(1,1)} B_\alpha^t \right)^2 \\ &+ \left( \sum_{|\alpha|=3} c_\alpha^t D_\xi^{(0,2)} B_\alpha^t \right)^2 \end{aligned} \right), \tag{3.89}
 \end{aligned}$$

(die rechte Seite der letzten Zeile ist eine Summe), und damit fortgesetzt

$$\begin{aligned}
 (3.89) &= \int_t \left( \left( c^{t\top} D_\xi^{(2,0)} \mathbf{B}^t \right)^2 + 2 \left( c^{t\top} D_\xi^{(1,1)} \mathbf{B}^t \right)^2 + \left( c^{t\top} D_\xi^{(0,2)} \mathbf{B}^t \right)^2 \right) \\
 &= \int_t c^{t\top} (\mathbf{B}_{\xi\xi}^t)^\top (c^{t\top} (\mathbf{B}_{\xi\xi}^t)^\top)^\top \\
 &= \int_t c^{t\top} ((\mathbf{B}_{\xi\xi}^t)^\top \mathbf{B}_{\xi\xi}^t) c^t. \tag{3.90}
 \end{aligned}$$

Da der Koeffizientenvektor  $c^t$  konstant ist, lassen sich die inneren Produkte mit ihm aus dem Integral ziehen und die letzte Zeile obiger Gleichung wird zu

$$\int_t c^{t\top} ((\mathbf{B}_{\xi\xi}^t)^\top \mathbf{B}_{\xi\xi}^t) c^t = c^{t\top} \left( \int_t (\mathbf{B}_{\xi\xi}^t)^\top \mathbf{B}_{\xi\xi}^t \right) c^t. \tag{3.91}$$

Das letzte Integral enthält per Konstruktion nur Polynome als Integranden und kann als solches über eine Quadraturformel exakt gelöst werden. Gültige Quadraturformeln solcher Art sind relativ leicht zu finden. Beispielsweise geben Zhang, T. Cui und Liu ein Verfahren an, Quadraturregeln zweiter (und höherer) Ordnung für baryzentrische Integrale auf Dreiecken zu konstruieren (Zhang, T. Cui und Liu, 2009). Fasst man  $\mathbf{B}_{\xi\xi}^t$  als Funktion von  $\boldsymbol{\lambda}$  auf, so ist beispielsweise die direkte Lösung  $E^t$  mit

$$E^t := \int_t (\mathbf{B}_{\xi\xi}^t)^\top \mathbf{B}_{\xi\xi}^t = \frac{|t|}{3} \left( \begin{aligned} &((\mathbf{B}_{\xi\xi}^t)^\top \mathbf{B}_{\xi\xi}^t)(0, \frac{1}{2}, \frac{1}{2}) \\ &+ ((\mathbf{B}_{\xi\xi}^t)^\top \mathbf{B}_{\xi\xi}^t)(\frac{1}{2}, 0, \frac{1}{2}) \\ &+ ((\mathbf{B}_{\xi\xi}^t)^\top \mathbf{B}_{\xi\xi}^t)(\frac{1}{2}, \frac{1}{2}, 0) \end{aligned} \right) \tag{3.92}$$

exakt, wobei  $|t|$  die Fläche (bzgl. den virtuellen Bildkoordinaten) des Dreiecks  $t$  ist. Nach dieser Konstruktion sind die  $E^t$  jeweils  $(10 \times 10)$ -Matrizen, die sich in einer quadratischen Zielfunktion des gesuchten Biegeenergie-Minimierungsproblems wiederfinden.

### 3.7.8 Quadratisches Minimierungsproblem

In den vorangegangenen Abschnitten 3.7.4 bis 3.7.7 wurden die Übergangs-Interpolations- und Energieminimierungsbedingungen für jeweils ein bzw. zwei Dreiecke als lineare Gleichungssysteme in den Koeffizienten  $c_\alpha^t$  der einzelnen Dreiecke aufgestellt. Um die gemachten Konstruktionen konsistent auf die gesamte Triangulation  $\Delta$  fortzusetzen, wird hier eine Methode konstruiert, alle auftretenden Gleichungssysteme der Einzeldreiecke zu koppeln.

Genau wird dasjenige Gleichungssystem gesucht, das als Argument einen Koeffizientenvektor  $c^\Delta$  besitzt, der jeden in der gesamten Triangulation vorkommenden Kontrollpunkt nur ein einziges Mal enthält, also insbesondere nur die nicht-doppelten Punkte (oder drei-, vier-, mehrfach an Knoten  $\xi$ , an denen mehrere Dreiecke zusammenstoßen). Beispielsweise lässt sich in den Abbildungen 3.19 und 3.20 ablesen, dass es bereits im Fall einer Triangulation aus zwei Dreiecken genau 16 einzigartige Kontrollpunkte gibt, obwohl beide Dreiecke für sich jeweils zehn besitzen<sup>11</sup>.

Der globale Kontrollnetzvektor  $c^\Delta$  darf für jeden (mehrfach-) redundanten Kontrollpunkt nur einen Eintrag enthalten. In welcher Reihenfolge dabei die nicht-redundanten Kontrollpunkte der jeweiligen Dreiecke referenziert werden, ist beliebig mit folgender Definition

**Definition 3.10 (Globaler Kontrollnetz-Ordinatenvektor)** Sei  $\Delta$  eine Triangulation der Ecken  $\xi$  gemäß Def. 3.7b und  $N_t = |\Delta|$  die Anzahl der Dreiecke  $t \in \Delta$ . Zu jedem Dreieck  $t$  sei der Kontrollnetz-Ordinatenvektor  $c^t$  der Länge 10 gegeben. Bezeichne  $N_c$  die Anzahl der einzigartigen Kontrollpunkte der gesamten Triangulation, dann wird ein  $N_c$ -Vektor  $c^\Delta$  Kontrollnetzvektor der Triangulation genannt, wenn für eine Menge

$$P_t \in \{0, 1\}^{10 \times N_c}$$

<sup>11</sup> Es sind dort die Punkte  $c_{10}^L$  und  $c_{10}^R$ ,  $c_8^L$  und  $c_9^R$ ,  $c_5^L$  und  $c_7^R$ , sowie  $c_1^L$  und  $c_4^R$  redundant.

### 3.7.8 Quadratisches Minimierungsproblem

von Permutationsmatrizen gilt

$$P_t c^\Delta = c^t. \quad (3.93)$$

Der Zweck dieser Definition wird deutlich, wenn man das Beispiel der  $\mathcal{C}^1$ -Bedingungen betrachtet: Zu einem beliebigen Fall „Lx-Ry“ aus der Tabelle 3.3 ist die Lineare Gleichung wie in (3.68) zu zwei Dreiecken  $t$  und  $t'$

$$C^{\text{Lx}} c^t = C^{\text{Ry}} c^{t'} \quad (3.94)$$

äquivalent zu einer homogenen Gleichung bezüglich des globalen Kontrollvektors:

$$\begin{aligned} C^{\text{Lx}} c^t &= C^{\text{Ry}} c^{t'} \\ \Leftrightarrow C^{\text{Lx}} P_t P_t^\text{T} c^t &= C^{\text{Ry}} P_{t'} P_{t'}^\text{T} c^{t'} \\ \Leftrightarrow C^{\text{Lx}} P_t c^\Delta &= C^{\text{Ry}} P_{t'} c^\Delta \\ \Leftrightarrow (C^{\text{Lx}} P_t - C^{\text{Ry}} P_{t'}) c^\Delta &= 0 \end{aligned} \quad (3.95)$$

Die Matrizen  $B_\xi^t$  der Gleichungen (3.81) bis (3.83) und die Matrizen  $E^t$  aus Gleichung (3.92) lassen sich auf die selbe Weise mit den Permutationen  $P_t$  auf den Vektor  $c^\Delta$  beziehen, wodurch die Systemmatrizen  $E$  (für „Energie“),  $S$  (für „Smoothness“),  $I$  (für „Interpolation“) und der Vektor  $G$  (für „Gradienten“) für das Gesamtproblem definiert werden können:

**Definition 3.11 (Systemmatrizen der Spline-Interpolation)** Es sei  $N_c$  die Anzahl der einzigartigen Kontrollpunkte der Triangulation  $\Delta$ ,  $N_i$  die Anzahl ihrer inneren Kanten,  $N_p$  die Anzahl ihrer Ecken,

$$G^\kappa := (G_1^{\kappa^\text{T}}, \dots, G_{N_p}^{\kappa^\text{T}})^\text{T}$$

der Vektor der Gradientennebenbedingungen zu irgendeiner Konfiguration  $\kappa$ , und  $P_t$  die Permutationsmatrizen nach (3.93) zu den Dreiecken  $t$ .

Dann sind mit dem oben beschriebenen Vektor  $c^\Delta$  der Länge  $N_c$  die System-Biegeenergiematrix

$$E \in \mathbb{R}^{N_c \times N_c} : \quad E := \sum_{t \in \Delta} P_t^\text{T} E^t P_t \quad (3.96)$$

mit  $E^t$  aus (3.92) definiert, die Matrix  $S$  der  $\mathcal{C}^1$ -Bedingungen des Splines

$$S \in \mathbb{R}^{3N_i \times N_c} : \quad S := \begin{pmatrix} & \vdots & \\ C^{\text{Lx}m} P_{t_m} - C^{\text{Ry}m} P_{t'_m} & & \\ & \vdots & \end{pmatrix}, \quad (3.97)$$

wobei die beiden Dreiecke  $t$  und  $t'$  die  $m$ -te innere Kante der Triangulation gemeinsam haben und  $C^{\text{Lx}m}$  nach (3.66) und  $C^{\text{Ry}m}$  nach (3.67) die dazu passenden Indexmatrizen sind, und die Matrix  $I$  der Interpolations-Nebenbedingungen

$$I \in \mathbb{R}^{2N_p \times N_c} : \quad I := \begin{pmatrix} & \vdots & \\ B_{\xi}^{t_m}(\lambda^{t_m}(\xi_m)) P_{t_m} & & \\ & \vdots & \end{pmatrix}, \quad (3.98)$$

für  $t_m$  jeweils ein beliebiges Dreieck, das  $\xi_m$  als Ecke enthält und  $B_{\xi}^{t_m}$  nach (3.77). Die  $m$ -ten zwei Gleichungen in  $I$  korrespondieren zum  $m$ -ten Gradienten in  $G^{\kappa}$ .

Die Konfigurationen der Systemmatrizen in obiger Definition hängen davon ab, in welcher Reihenfolge die Punkte  $\xi_i$  geordnet sind, die Triangulation  $\Delta$  über sie gebildet wird, für welches angrenzende Dreieck die Interpolationsbedingung an jedem Punkt  $\xi_i$  ausgewertet wird, in welcher Richtung die Übergangsbedingungen gebildet werden, etc. Alle möglichen Matrizen beschreiben aber zu einem gegebenen Kontrollpunktvektor  $c^{\Delta}$  denselben Spline. Die  $\mathcal{C}^0$ -Übergangsbedingungen sind oben nicht aufgeführt. Das liegt daran, dass sie hier durch die Eliminierung redundanter Kontrollpunkte trivialerweise erfüllt sind.

Auf Seite 75 wurde das zu lösende Interpolationsproblem in Worten formuliert. Mit den in Definition 3.11 eingeführten Systemmatrizen lässt sich die Formulierung präzisieren:

Die gesuchte Oberflächenfunktion  $\hat{r}(\xi)$  ist Exponential eines Splines  $s(\xi)$  der Art (3.64) aus Definition 3.9, der bestimmt ist über einen Vektor  $c^\Delta$  von Kontrollpunkten.  $c^\Delta$  ist eindeutige Lösung der quadratischen Minimierung mit linearer Gleichheits-Nebenbedingung

$$\begin{aligned} \min_c \quad & \frac{1}{2} c^\top E c \\ \text{unter} \quad & \begin{pmatrix} 1, 0, \dots, 0 \\ S \\ I \end{pmatrix} c = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ G^\kappa \end{pmatrix} \end{aligned} \quad (3.99)$$

mit Matrizen  $E$ ,  $S$ ,  $I$  und  $G^\kappa$  aus Definition 3.11.

Die Größen der Nebenbedingungen lassen sich aus den Gleichungen (3.96), (3.97) und (3.98) ablesen. Die erste Zeile der Nebenbedingungen legt den letzten Freiheitsgrad des Problems, die Entfernung des ersten Kontrollpunktes der Splinefläche vom Kamera-Ursprung, auf  $\exp(0) = 1$  fest.

Die Konstruktion der Systemmatrizen aus Definition 3.11 kann in vollem Detail anhand des einfachen Beispiels A.3 in Anhang A für ein Problem mit vier Stützstellen nachvollzogen werden.

### 3.7.9 Lösungsverfahren für das Minimierungsproblem

Die LAGRANGE-Funktion für das Minimierungsproblem (3.99) ist

$$\mathcal{L}(c, l_1, l_2, l_3) = \frac{1}{2} c^\top E c + l_1 c_1 + l_2^\top S c + l_3^\top (I c - G^\kappa), \quad (3.100)$$

mit LAGRANGE-Multiplikatoren  $l_1$  bis  $l_3$ . Die Extremstellen von  $\mathcal{L}$  sind dann die Kandidaten für die optimalen Punkte des ursprünglichen Minimierungsproblems. Die Ableitung von (3.100) nach  $c$  und den  $l_i$  ist

$$\begin{aligned} E c + l_1 + S l_2^\top + I l_3^\top &= 0 \\ c_1 &= 0 \\ S c &= 0 \\ I c &= G^\kappa \end{aligned} \quad (3.101)$$

beziehungsweise als in den  $l_i$  und  $c$  lineares System, mit  $e_1$  dem ersten kanonischen Einheitsvektor des  $\mathbb{R}^{N_c}$ ,

$$\begin{pmatrix} E & e_1 & S^T & I^T \\ e_1^T & & & \\ S & & 0 & \\ I & & & \end{pmatrix} \begin{pmatrix} c \\ l_1 \\ l_2 \\ l_3 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ G^\kappa \end{pmatrix}. \quad (3.102)$$

Für dieses System kann über gängige Verfahren eine direkte Lösung ermittelt werden. Als Alternative zeigten Awanou, Lai und Wenston für ein sehr ähnliches Interpolationsproblem, dass die folgende iterative Lösung gegen den gesuchten Kontrollnetz-Ordinatenvektor  $c^\Delta$  konvergiert: Definiere

$$B := (e_1, S^T, I^T)^T, \quad g^\kappa := (0^{1 \times (3N_i+1)}, (G^\kappa)^T)^T \quad (3.103)$$

und  $\varepsilon$  hinreichend klein, setze

$$c^{(0)} := \left( E + \frac{1}{\varepsilon} B^T B \right)^{-1} \left( \frac{1}{\varepsilon} B^T g^\kappa \right), \quad (3.104)$$

und iteriere

$$c^{(k+1)} := \left( E + \frac{1}{\varepsilon} B^T B \right)^{-1} \left( E c^{(k)} + \frac{1}{\varepsilon} B^T g^\kappa \right) \quad (3.105)$$

bis eine Toleranzschwelle bezüglich  $\|c^{(k+1)} - c^{(k)}\|$  unterschritten ist (Awanou, Lai und Wenston, 2005). Diese Methode hat den erheblichen Vorteil, dass alle Matrizen maximale Größe  $N_c$  besitzen, während das Problem (3.102) erheblich größer sein kann.

$\alpha$ -#	$D_{\xi}^{(1,0)} B_{\alpha}^t$	$D_{\xi}^{(0,1)} B_{\alpha}^t$
1	$-\frac{3}{e} \lambda_3^2(a+c)$	$-\frac{3}{e} \lambda_3^2(b+d)$
2	$\frac{3}{e} (\lambda_3^2 c - 2\lambda_2 \lambda_3(a+c))$	$\frac{3}{e} (\lambda_3^2 d - 2\lambda_3 \lambda_2(b+d))$
3	$\frac{3}{e} (2\lambda_3 \lambda_2 c - \lambda_2^2(a+c))$	$\frac{3}{e} (2\lambda_3 \lambda_2 d - \lambda_2^2(b+d))$
4	$\frac{3}{e} \lambda_2^2 c$	$\frac{3}{e} \lambda_2^2 d$
5	$\frac{3}{e} (\lambda_3^2 a - 2\lambda_3 \lambda_1(a+c))$	$\frac{3}{e} (\lambda_3^2 b - 2\lambda_3 \lambda_1(b+d))$
6	$\frac{3}{e} 2(\lambda_3 \lambda_2 a + \lambda_3 \lambda_1 c - \lambda_2 \lambda_1(a+c))$	$\frac{3}{e} 2(\lambda_3 \lambda_2 b + \lambda_3 \lambda_1 d - \lambda_2 \lambda_1(b+d))$
7	$\frac{3}{e} (\lambda_2^2 a + 2\lambda_2 \lambda_1 c)$	$\frac{3}{e} (\lambda_2^2 b + 2\lambda_2 \lambda_1 d)$
8	$\frac{3}{e} (2\lambda_3 \lambda_1 a - \lambda_1^2(a+c))$	$\frac{3}{e} (2\lambda_3 \lambda_1 b - \lambda_1^2(b+d))$
9	$\frac{3}{e} (2\lambda_2 \lambda_1 a + \lambda_1^2 c)$	$\frac{3}{e} (2\lambda_2 \lambda_1 b + \lambda_1^2 d)$
10	$\frac{3}{e} \lambda_1^2 a$	$\frac{3}{e} \lambda_1^2 b$

Tabelle 3.4: Ableitungen von  $B_{\alpha}^t$  nach den virtuellen Bildkoordinaten für alle  $\alpha$



$\alpha$ -#	$D_{\xi}^{(2,0)} B_{\alpha}^t$	$D_{\xi}^{(0,2)} B_{\alpha}^t$
1	$\frac{6}{e^2} \lambda_3(a+c)^2$	$\frac{6}{e^2} \lambda_3(b+d)^2$
2	$\frac{6}{e^2} (\lambda_2(a+c) - 2\lambda_3c)(a+c)$	$\frac{6}{e^2} (\lambda_2(b+d) - 2\lambda_3d)(b+d)$
3	$\frac{6}{e^2} (\lambda_3c^2 - 2\lambda_2c(a+c))$	$\frac{6}{e^2} (\lambda_3d^2 - 2\lambda_2d(b+d))$
4	$\frac{6}{e^2} \lambda_2c^2$	$\frac{6}{e^2} \lambda_2d^2$
5	$\frac{6}{e^2} (\lambda_1(a+c) - 2\lambda_3a)(a+c)$	$\frac{6}{e^2} (\lambda_1(b+d) - 2\lambda_3b)(b+d)$
6	$\frac{6}{e^2} 2(\lambda_3ac - (\lambda_2a(a+c) + \lambda_1c(a+c)))$	$\frac{6}{e^2} 2(\lambda_3bd - (\lambda_2b(b+d) + \lambda_1d(b+d)))$
7	$\frac{6}{e^2} (2\lambda_2ac + \lambda_1c^2)$	$\frac{6}{e^2} 2\lambda_2bd + \lambda_1d^2$
8	$\frac{6}{e^2} (\lambda_3a^2 - 2\lambda_1a(a+c))$	$\frac{6}{e^2} (\lambda_3b^2 - 2\lambda_1b(b+d))$
9	$\frac{6}{e^2} (\lambda_2a^2 + 2\lambda_1ac)$	$\frac{6}{e^2} (\lambda_2b^2 + 2\lambda_1bd)$
10	$\frac{6}{e^2} \lambda_1a^2$	$\frac{6}{e^2} \lambda_1b^2$
$\alpha$ -#	$D_{\xi}^{(1,1)} B_{\alpha}^t$	
1	$\frac{6}{e^2} \lambda_3(a+c)(b+d)$	
2	$\frac{6}{e^2} (\lambda_2(a+c)(b+d) - \lambda_3(c(b+d) + d(a+c)))$	
3	$\frac{6}{e^2} (\lambda_3cd - \lambda_2(c(b+d) + d(a+c)))$	
4	$\frac{6}{e^2} \lambda_2cd$	
5	$\frac{6}{e^2} (\lambda_1(a+c)(b+d) - \lambda_3(a(b+d) + b(a+c)))$	
6	$\frac{6}{e^2} (\lambda_3(ad+bc) - \lambda_2(a(b+d) + b(a+c)) - \lambda_1(c(b+d) + d(a+c)))$	
7	$\frac{6}{e^2} (\lambda_2(ad+bc) + \lambda_1cd)$	
8	$\frac{6}{e^2} (\lambda_3ab - \lambda_1(a(b+d) + b(a+c)))$	
9	$\frac{6}{e^2} (\lambda_2ab + \lambda_1(ad+bc))$	
10	$\frac{6}{e^2} \lambda_1ab$	

 Tabelle 3.5: Zweite partielle Ableitungen von  $B_{\alpha}^t$  nach den virtuellen Bildkoordinaten für alle  $\alpha$

### 3.7.10 Zwischenfazit

In den vorangehenden Unterabschnitten wurden gezeigt, wie die abstrakten Definitionen zur Datenmenge, Konfiguration, Oberflächenbeschaffenheit, etc., auf die Konstruktion eines eindeutigen linearen Gleichungssystems führen, dessen Lösung die gesuchte Oberfläche approximiert. Währenddessen blieb offen, auf welche Art die Konfiguration  $\kappa$  gewonnen werden kann, die Parameter des Gleichungssystems ist. Dies wird im folgenden Abschnitt behandelt, wo aus der Biegeenergie-Eigenschaft der Oberflächenrekonstruktion eine Art schwache Eindeutigkeit für die „richtige“ Konfiguration hergeleitet wird. Ein Algorithmus zur Herstellung der Konfiguration wird dort direkt angegeben.

Zweitens bleibt bis zu diesem Punkt noch unklar, inwieweit die Verwendung der mittleren virtuellen Kraterkoordinaten  $\xi$  aus Def. 3.6 anstatt der zur tatsächlichen Projektion der Kraterzentren gehörenden virtuellen Bildkoordinaten Fehler in den Lösungskoeffizienten der Oberflächenrekonstruktion erzeugt. Dies wird im Kapitel 4, Abschnitt 4.3, untersucht.

## 3.8 Konfigurationsalgorithmus

### 3.8.1 Einleitung

Das in Abschnitt 3.7 vorgestellte Verfahren ist geeignet, zu einem beliebigen Satz von Daten nach Definition 3.1a und einer *fest gewählten* Konfiguration dieser Daten nach Definition 3.1b eine biegeenergie-minimale Ausgleichsfläche aus Splines zu erzeugen, auf die die erkannten Krater projiziert werden können.

Bisher wurde die konkrete Auswahl der Konfiguration  $\kappa$  vernachlässigt. Es existiert eine Menge von  $2^{N_p}$  möglichen Konfigurationen bei  $N_p$  erkannten Kratern, unter ihnen die „richtige“. Darunter versteht sich in diesem Zusammenhang genau diejenige Konfiguration, die alle Normalenvektoren derart orientiert, dass aus ihnen mit dem vorgestellten Verfahren die getreueste Approximation an die im Bild beobachtete physische Oberfläche rekonstruiert wird. Die in diesem Sinne optimale Konfiguration zu isolieren, ist Gegenstand dieses Abschnittes.

### 3.8.2 Iterative energieminimale Konfiguration

Jedem erkannten Krater im Bild sind zwei Tupel von je einem Zentren- und einem Normalenvektoren  $\mathbf{r} = (c, n)$  und  $\mathbf{r}' = (c', n')$  aus der Gesamtmenge der

Erkennungsdaten  $\mathbf{R}$  zugeordnet, und daraus abgeleitet eine virtuelle Bildkoordinate  $\xi$  gemäß Def. 3.6. Jedes beliebige Dreieck  $t \in \Delta$  aus der Triangulation der Punkte  $\xi$  referenziert nach Def. 3.7a drei Datentupel  $\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k$ , die auf  $2^3 = 8$  verschiedene Arten konfiguriert werden können.

Die Grundlage zum folgenden Konfigurationsansatz bildet die Annahme, dass aus allen möglichen Oberflächen, die sich aus den unkonfigurierten Daten  $\mathbf{R}$  rekonstruieren lassen, diejenige mit der minimalen Biegeenergie die korrekt konfigurierte repräsentiert. Aus dieser *globalen* Annahme über die auf ganz  $\Delta$  definierten Oberfläche lässt sich für einzelne  $t \in \Delta$  ableiten, dass auch die Oberfläche zur *lokalen* Konfiguration der an  $t$  beteiligten Krater energieminimal sein sollte. Da zu jedem Dreieck  $t$  nur 8 verschiedene Konfigurationen existieren, ist es zu jedem einzelnen Dreieck möglich, sie alle bezüglich ihrer Biegeenergie zu vergleichen und die energieminimale unter ihnen auszuwählen.

Zu einem beliebigen Dreieck  $t_0$  aus einer gegebenen Triangulation  $\Delta$  liefert die Funktion „`direct_configuration_single()`“ aus Algorithmus 2 einen solchen energieminimalen Konfigurationskern: Es wird zur von  $t_0$  dargestellten Teiltriangulation das Splinesystem aufgestellt (vgl. Funktion „`config_to_coeffs()`“ aus Algorithmus 1 mit Referenzen) und zu jeder möglichen Konfiguration der Biegeenergie-Term verglichen. Die Konfiguration an dessen Minimum ist die gesuchte Startkonfiguration.

Diese Konfiguration der Datenpunkte des ersten Dreiecks ist der Startpunkt zur Funktion „`iterative_configuration()`“, der iterativen Konfiguration der gesamten Datenmenge, die in Algorithmus 5 formal in Pseudocode beschrieben ist: Ausgehend vom Startdreieck lässt sich ein weiteres Dreieck aus  $\Delta$  ohne  $t_0$  hinzufügen. Dazu durchsucht man wie in der Funktion „`next_triangle()`“ aus Algo. 3 diese verbleibende Differenztriangulation nach einem an die bisher konfigurierte Triangulation angrenzenden neuen Dreieck  $t$ . In den meisten Fällen kommt damit auch ein weiterer Datenpunkt aus  $\mathbf{R}$  hinzu, dessen Index man durch Funktion „`next_point_idx()`“ aus Algo. 4 findet, und damit zwei Möglichkeiten, die resultierende Menge von jetzt vier Datenpunkten zu konfigurieren (da die ersten drei als bereits fest konfiguriert betrachtet werden). Der Vergleich der beiden möglichen neuen Konfigurationen mit Hinblick auf die Biegeenergie der resultierenden Fläche liefert dann die Entscheidung für eine der beiden Möglichkeiten. Auf diese Weise lässt sich mit linearem Aufwand  $\mathcal{O}(N_p)$  die gesamte Datenmenge  $\mathbf{R}$  durchkonfigurieren, im Gegensatz zum exponentiellen Aufwand  $\mathcal{O}(2^{N_p})$  der naiven direkten Suche durch alle möglichen Konfigurationen.

Dieses Vorgehen bei der iterativen Konfiguration wird in Tabelle 3.6 am Beispiel von neun regulär angeordneten Punkten auf einer Fläche vor der

### 3.8.2 Iterative energieminimale Konfiguration

Kamera illustriert. In den Abbildungen sind auf dem im jeweiligen Schritt bereits konfigurierten Teil der Fläche die Radius-Isolinien abgetragen, d.h. die Schnitte der Fläche mit Kugeln von konstantem Radius um den Kamerakoor-  
dinaten-Ursprung, und dazu der  $(\nu, \varphi)$ -Gradient zum aktuell hinzugefügten und zu konfigurierenden Punkt.

Tabelle 3.6: Isolinien und Gradienten der rekonstruierten Oberflächen während der iterativen Konfiguration der Daten **R**

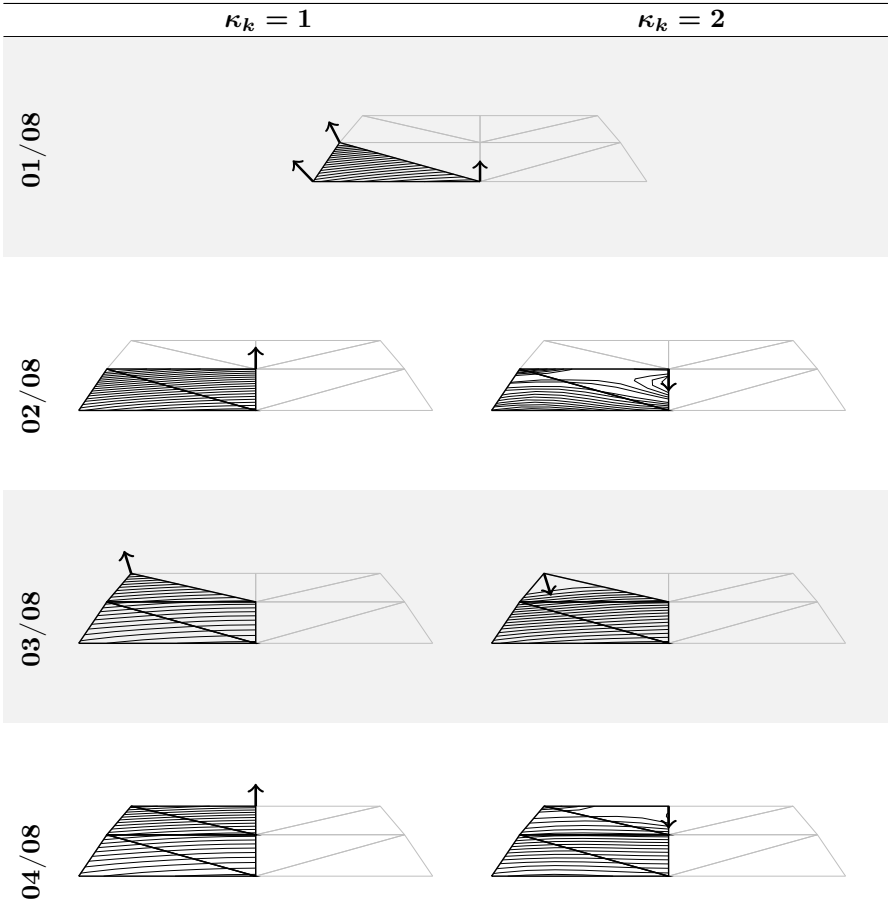
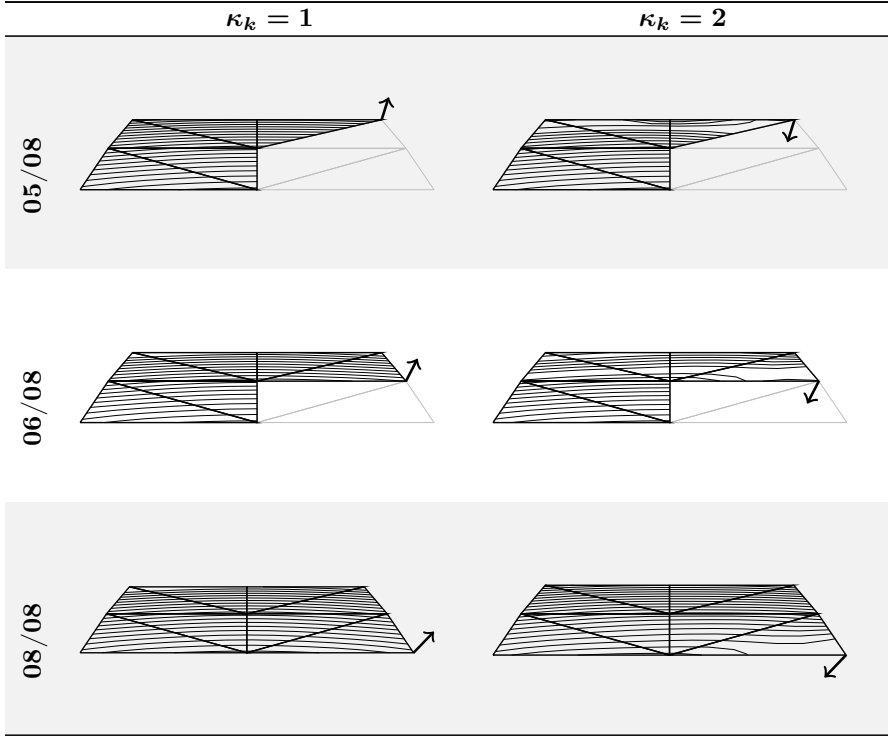


Tabelle wird fortgesetzt auf folgender Seite

Fortsetzung Tabelle 3.6



In der ersten Zeile der Tabelle 3.6 ist die Fläche für die durch Algorithmus 2 gefundene Startkonfiguration gezeigt. In den Zeilen darunter ist jeweils der Vergleich der aus dem Hinzufügen eines weiteren Datenpunktes resultierenden zwei Konfigurationsmöglichkeiten gegenübergestellt. Durch Vergleich der Glattheit und Regelmäßigkeit der Isolinien lässt sich sofort die Konfiguration mit kleinerer resultierender Biegeenergie ablesen.

Der Schritt zu Dreieck 07/08 (mit Bezug auf die linke Spalte in der Tabelle) ist nicht gezeigt, da das Hinzufügen keinen weiteren Datenpunkt und damit keine Konfigurationsmöglichkeit hinzufügt. Man erkennt hieran, dass der Konfigurationsaufwand von der Anzahl der Punkte  $N_p$  und nicht von der Anzahl der Dreiecke  $N_t$  abhängt (die aber auch nur gegen  $2N_p$  geht für  $N_p$  groß).

Die Algorithmen 1 bis 5 basieren auf der Annahme, dass der Konfigurationskern zum Startdreieck energieminimal ist und daher gültiger Teil der

```

1 Function config_to_coeffs( $\mathbf{R}, \Delta, \kappa$ )
2    $A \leftarrow \begin{pmatrix} E & e_1 & S^T & I^T \\ e_1^T & & & \\ S & & 0 & \\ I & & & \end{pmatrix}_{\mathbf{R}, \Delta}$  /* Def. 3.11 */
3    $b \leftarrow (0, \dots, G^\kappa)_{\mathbf{R}, \Delta}^T$  /* Glchg. (3.78) */
4    $x \leftarrow A^{-1}b$ 
5   return  $x_{(1, \dots, N_c)}$  /* Erste  $N_c$  Elemente sind  $c^\Delta$ . */

```

**Algorithmus 1 :** Hilfsfunktion zur Berechnung der Spline-Koeffizienten aus einer konfigurierten Teiltriangulation.

```

1 Function direct_configuration_single( $\mathbf{R}, t$ )
2    $e_{\min} \leftarrow \infty$ 
3    $\kappa_{\text{opt}} \leftarrow \emptyset$ 
4   for all  $\kappa \in \{1, 2\}^3$  do
5      $c^t \leftarrow \text{config\_to\_coeffs}(\mathbf{R}_t, t, \kappa)$  /* Algo. 1. */
6      $e \leftarrow (c^t)^T (E_{\mathbf{R}_t, t}) c^t$ 
7     if  $e_{\min} > e$  then
8        $e_{\min} \leftarrow e$ 
9        $\kappa_{\text{opt}} \leftarrow \kappa$ 
10  return  $\kappa_{\text{opt}}$ 

```

**Algorithmus 2 :** Hilfsfunktion zum Ermitteln der energieminimalen Konfiguration eines einzelnen Dreiecks.

```

1 Function next_triangle( $\Delta, \Delta^C$ )
2   for  $t \in \Delta \setminus \Delta^C$  do
3     if  $(t \cup \Delta^C)$  ist Teiltriangulation von  $\Delta$  /* vgl. Def. 3.7c */
4     then
5       return  $t$ 
6   return  $\emptyset$ 

```

**Algorithmus 3 :** Hilfsfunktion zum finden des nächsten Dreiecks während der iterativen Konfiguration.

```

1 Function next_point_idx( $\Delta, t$ )
2   for  $i = 1, 2, 3$  do
3     if  $\nexists t' \in \Delta : (t_i = t'_1 \text{ or } t_i = t'_2 \text{ or } t_i = t'_3)$  then
4       return  $t_i$ 
5 return  $\emptyset$ 

```

**Algorithmus 4 :** Hilfsfunktion zum finden des Index eines Punktes eines neuen Dreiecks, der noch nicht von der Triangulation abgedeckt war.

```

1 Function iterative_configuration( $\mathbf{R}, \Delta$ )
2   wähle  $t_0$  beliebig aus  $\Delta$ 
3    $\kappa \leftarrow \text{direct\_configuration\_single}(\mathbf{R}, t_0)$       /* Algo. 2. */
4    $\Delta^C \leftarrow t_0$ 
5   while  $\Delta^C \neq \Delta$  do
6      $e_{\text{opt}} \leftarrow \infty$ 
7      $t \leftarrow \text{next\_triangle}(\Delta, \Delta^C)$               /* Algo. 3 */
8      $i \leftarrow \text{next\_point\_idx}(\Delta^C, t)$            /* Algo. 4 */
9      $\mathcal{I} \leftarrow (\text{idx}(\Delta^C), i)$                   /* Def. 3.7d */
10    for  $k \in \{1, 2\}$  do
11       $c \leftarrow \text{config\_to\_coeffs}(\mathbf{R}_{\mathcal{I}}, \Delta^C \cup t, (\kappa, k))$  /* Algo. 1 */
12       $e \leftarrow c^T(E_{\mathbf{R}_{\mathcal{I}}, \Delta^C \cup t})c$ 
13      if  $e < e_{\text{opt}}$  then
14         $e_{\text{opt}} \leftarrow e$ 
15         $\kappa_{\text{opt}} \leftarrow (\kappa, k)$ 
16     $\kappa \leftarrow \kappa_{\text{opt}}$ 
17     $\Delta^C \leftarrow \Delta^C \cup t$ 
18  return  $\kappa$ 

```

**Algorithmus 5 :** Energieminimale iterative Konfiguration der Triangulation

### 3.8.2 Iterative energieminimale Konfiguration

---

optimalen Endkonfiguration. Das ist leider nicht uneingeschränkt der Fall. Die Definition 3.1c führte den Begriff der *symmetrischen Konfiguration* ein. Zur Erinnerung: Damit ist die Konfiguration  $\kappa'$  bezeichnet, die zu einer gegebenen Konfiguration  $\kappa$  komplementär ist in dem Sinne, dass  $\kappa'$  überall dort 1 ist, wo  $\kappa$  2 ist, und umgekehrt.

Das praktische Problem in diesem Zusammenhang ist, dass es zu einem Startdreieck zur oben vorgestellten iterativen Konfiguration nicht nur eine energieminimale Konfiguration gibt, sondern zwei gleicher Biegeenergie. Je nachdem, welche der beiden Startkonfigurationen gewählt wird, erreicht man über die iterative Konfiguration einen Endzustand, der symmetrisch im Sinne der Def. 3.1c zu dem Endzustand ist, den man ausgehend von der anderen Startkonfiguration erreicht hätte. Nur eine der beiden Endkonfigurationen kann aber eine rekonstruierte Oberfläche erzeugen, die eine nützliche Approximation an die beobachtete Oberfläche ist und sich daher eignet, die Tiefenunsicherheit der beobachteten Krater aufzulösen.

Es zeigt sich glücklicherweise, dass diese unentscheidbare Symmetrie kein echtes Hindernis ist. Die folgende Behauptung liefert dazu zwei wichtige Eigenschaften der Oberflächen zu symmetrisch konfigurierten Daten:

#### **Behauptung 3.3 (Biegeenergie symmetrisch konfigurierter Daten)**

Es seien  $\mathbf{R}$  rekonstruierte Zentren- und Normalenvektoren-Daten zu einer Menge von erkannten Kratern nach Def. 3.1a,  $\kappa$  eine beliebige Konfiguration der Daten nach Def. 3.1b und  $\Delta$  eine beliebige Triangulation nach Def. 3.7 der mittleren virtuellen Kraterzentren nach Def. 3.6.

Für die symmetrische Konfiguration  $\kappa'$  im Sinne der Definition 3.1c zu  $\kappa$  und den eindeutigen Kontrollnetz-Ordinatenvektor  $c^\Delta$ , der für  $\kappa$  aus der Anwendung des in Abschnitt 3.7 konstruierten Verfahrens resultiert, ergibt sich der zu  $\kappa'$  resultierende Vektor  $c^{\Delta'}$  als

$$c^{\Delta'} = -c^\Delta. \quad (3.106)$$

Die Biegeenergie der aus  $\mathbf{R}^{[\kappa']}$  rekonstruierten Oberfläche ist dieselbe wie die der aus  $\mathbf{R}^{[\kappa]}$  rekonstruierten.

**Beweis:** Der Beleg für diese Behauptung ergibt sich direkt aus der Konstruktion der beteiligten Systeme: Die Stützstellen des Interpolationsproblems wurden als die mittleren virtuellen Bildkoordinaten  $\xi$  der Kraterzentren angenommen, und nur dort werden die beteiligten Größen ausgewertet. Diese Koordinaten sind genau die sphärischen Koordinaten  $\nu$  und  $\varphi$  der Projektionskegel-Mittelstrahlen  $q$  aus der Rekonstruktion der räumlichen Kraterparameter



im Abschnitt 3.3. Die Konfigurationsalternativen  $n$  und  $n'$  für die Normalenvektoren standen mit diesen im Zusammenhang

$$n' = 2qq^T n - n. \quad (3.107)$$

Bezeichne  $J_q$  kurz die  $\nu$ - und  $\varphi$ -Spalten der JACOBI-Matrix der Transformation von sphärischen nach kartesischen Kamerakoordinaten ist an diesen Stellen  $(\nu_q, \varphi_q, 1) = T_C^\Phi(q)$

$$J_q := J[T_\Phi^C](\nu_q, \varphi_q) = \begin{pmatrix} \cos \nu_q \cos \varphi_q & -\sin \nu_q \sin \varphi_q \\ \cos \nu_q \sin \varphi_q & \sin \nu_q \cos \varphi_q \\ -\sin \nu_q & 0 \end{pmatrix}. \quad (3.108)$$

Wegen der Symmetrie (3.107) ist an den Stellen  $q$  der Vorfaktor  $1/q^T n = 1/q^T n'$ . Damit ist die Summe der skalierten  $\nu$ - $\varphi$ -Gradienten  $\nabla r/r$  aus Gleichung (3.54) an den Stellen  $(\nu_q, \varphi_q)$  für  $n$  und  $n'$ :

$$\begin{aligned} \frac{n^T}{q^T n} J_q + \frac{(n')^T}{q^T n'} J_q &= \frac{1}{q^T n} (n^T + (n')^T) J_q \\ &= \frac{1}{q^T n} (n^T + (2qq^T n)^T - n^T) J_q \\ &= \frac{1}{q^T n} ((2qq^T n)^T) J_q \\ &= \frac{1}{q^T n} 2n^T qq^T J_q \end{aligned} \quad (3.109)$$

Der Vektor  $q$  ist aber gerade die letzte Spalte von  $J[T_\Phi^C]$ , weswegen  $q^T J_q$  immer  $(0, 0)$  ist (Orthonormalität der Basis). Daher ist die gesamte rechte Seite

$$\frac{1}{q^T n} 2n^T \underbrace{q q^T J_q}_{(0,0)} = (0, 0) \quad (3.110)$$

und deswegen

$$\frac{n^T}{q^T n} J_q = -\frac{(n')^T}{q^T n'} J_q, \quad (3.111)$$

also  $G^{\kappa} = -G^{\kappa'}$ . In die Konstruktion der Matrizen  $E$ ,  $S$ ,  $I$  der Def. 3.11 gehen wieder nur die auf den Koordinaten  $\xi$  (bzw.  $q$ ) basierenden Koeffizienten  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$  aus Gleichung (3.69) ein. Beim Wechsel von  $\kappa$  auf  $\kappa'$  ändert sich im LAGRANGE-System aus Gleichung (3.102) deswegen nur das Vorzeichen der rechten Seite. Da das System per Konstruktion eindeutig lösbar ist, müssen

### 3.8.3 Zwischenfazit

---

die Lösungen der Systeme zu den beiden rechten Seiten bis auf Vorzeichen gleich sein. In die Biegeenergie-Gleichung geht diese Lösung im Quadrat ein, ihr Ergebnis ändert sich daher nicht.  $\square$

Es macht also keinen Unterschied, von welcher der beiden energieminimalen Startkonfigurationen man ausgeht. Erreicht man ausgehend von der „falschen“ Startkonfiguration über die Iteration einen energieminimalen Endzustand, der symmetrisch zu demjenigen ist, der die tatsächlich beobachtete Oberfläche bestmöglich approximieren würde, wechselt man schlicht das Vorzeichen der Lösung. Es ist so zwar nötig, die Korrespondenzbildung zum Kraterkatalog je zweimal pro Datensatz  $\mathbf{R}$  durchzuführen, einmal für die positiven und einmal für die negativen ermittelten Koeffizienten  $c^{\Delta}$ . Die Systemmatrizen der Spline-Interpolation müssen aber nur einmal aufgestellt werden.

Für den Biegeenergie-Vergleich im Algorithmus 5 lässt sich auf dieselbe Art Aufwand einsparen. Bei gleichbleibender Problemgröße muss beim Wechsel einer Komponente der Konfiguration nur das Vorzeichen der zwei korrespondierenden Gradienten-Komponenten gewechselt werden. Die LA-GRANGE-Sytemmatrizen müssen bei so einem Wechsel in der Hilfsfunktion „`config_to_coeffs()`“ nicht wieder neu konstruiert werden.

### 3.8.3 Zwischenfazit

Die Algorithmen 1 bis 5 liefern das Pseudocode-Gerüst für ein  $\mathcal{O}(N_p)$ -Verfahren zur vollständigen iterativen Konfiguration der Daten  $\mathbf{R}$ . Die Aussagen der belegten Behauptung 3.3 stellen darüber hinaus sicher, dass auch bei unentscheidbarer Startkonfiguration das gesuchte Ergebnis maximal einen Vorzeichenwechsel der resultierenden Kontrollnetz-Ordinaten erfordert.

Was damit vorliegt, ist eine Referenz-Splinefläche, mittels derer man die relativen Tiefeninformationen der beobachteten Krater auflösen kann. Zu jedem Tupel  $(\hat{c}_i, \hat{n}_i)$  eines Zentrums-Richtungsvektors und einer Oberflächennormale, das Komponente der so konfigurierten Daten  $\mathbf{R}^{[\kappa]}$  ist, ergeben sich damit die  $\lambda$  aus Definition 3.2 (nicht zu verwechseln mit den baryzentrischen Koordinaten) bis auf einen globalen Skalar  $\Lambda$  (vgl. Behauptung 3.2 in Abschnitt 3.6.2 und darauf folgende Erläuterungen). Die zu  $\hat{c}_i$  gehörende Tiefe  $\lambda_i$  erhält man durch Auswertung der rekonstruierten Spline-Oberfläche an der Stelle  $T_C^V(\hat{c}_i)$ . Lässt man die Approximation  $\hat{c}_i \approx q_i$  zu, so kann man  $\lambda_i$  für die an der Spline-Interpolation beteiligten Krater direkt aus dem Kontrollnetz-Ordinatenvektor ablesen, da jede Stützstelle, bzw. ihre korrespondierende virtuelle Bildkoordinate  $\xi$ , auch Kontrollpunkt und ihre Ordinate damit Kom-

ponente von  $c^\Delta$  ist. Für alle übrigen Krater, deren virtuelle Bildkoordinaten von der Splinefläche abgedeckt werden, muss der Spline einmal ausgewertet werden. Dies ist numerisch stabil und günstig über den Algorithmus von DE CASTELJAU möglich.

## 3.9 Korrespondenzen zum Katalog und Kamerateapose

### 3.9.1 Einleitung

Die Spline-Oberflächenrekonstruktion mit dem oben vorgestellten Konfigurationsverfahren für die Daten liefert alle Informationen, die für die Korrespondenzbildung zwischen den erkannten Kratern und dem mitgeführten Kraterkatalog notwendig sind:

- Abschnitt 3.3 beschreibt die Rekonstruktion der kamerafesten Richtungsvektoren zu den Zentren beobachteter Krater und die der kamerafesten Oberflächennormalen an diesen Punkten, sowie der Einheitsradien<sup>12</sup>  $\bar{\rho}$  der beobachteten Krater.
- In Abschnitt 3.7 wird ein Verfahren konstruiert, aus diesen kamerafesten Richtungsvektoren Ausgleichsflächen zu berechnen, die eine Tiefenauflösung  $\lambda$  für die beobachteten Krater liefern.
- Abschnitt 3.8 stellt eine Heuristik zur Verfügung, die aus allen möglichen Ausgleichsflächen eine (bzw. zwei) optimale über die Konfiguration  $\kappa$  isoliert.

Mit diesen Daten und zusammen mit einem gegebenen Katalog  $\mathbf{K}$  von bekannten mondfesten Kratermittelpunkten und -Radien erhält man ein klassisches Zuordnungsproblem. In Abb. 3.21 ist es im Gesamtzusammenhang mit dem Datenfluss zwischen den oben genannten Verfahrensstufen dargestellt. Zwischen den „Datenlisten“ sind tatsächliche (aber unbekannte) Korrespondenzen mit Pfeilen markiert. Dies sind die Zuordnungen, die algorithmisch gefunden werden sollen. Schwierigkeiten der o.g. Zuordnung sind die folgenden, teilweise im Diagramm der Abb. 3.21 zu erkennen:

---

<sup>12</sup>Radius des Kreisschnittes des Projektionskegels in Entfernung eins vom Kameraursprung.

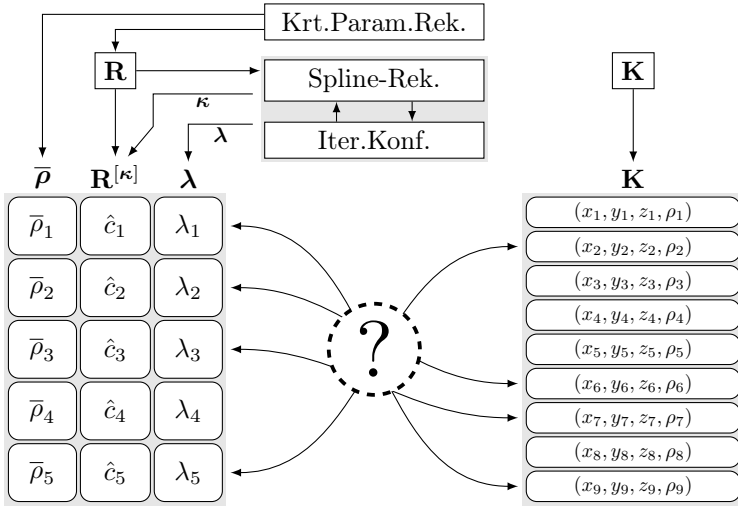


Abbildung 3.21: Beispieldiagramm Zuordnungsproblem Kraterkatalog zu rekonstruierten Daten der beobachteten Krater mit Datenfluss.

- Es gibt beobachtete Krater, zu denen keine Katalogeinträge vorliegen (keine Pfeile gehen von ihnen aus). Man kann sagen, dass unbekannte Krater im Bild erkannt werden.
- Es gibt Katalogeinträge, die nicht im Bild beobachtet wurden (es enden keine Pfeile an ihnen). Das kann vielerlei Gründe haben, von Szenarioparametern wie Beleuchtungsbedingungen über algorithmischen Gegebenheiten bis hin zur Möglichkeit, dass aus der aktuellen Position und Lage heraus Teile der bekannten Krater im Katalog schlicht nicht sichtbar sind.
- Die Datenlisten der Beobachtungen sind in keiner vorhersagbaren Weise geordnet.
- Die Einträge des Kataloges sind metrisch auf Weltskala (üblicherweise MCMF (Moon-Centered Moon-Fixed)), die Beobachtungsdaten sind mit einer unbekannten Größe  $\Lambda$  skaliert (vgl. Beh. 3.2 aus Abschn. 3.6.2), also praktisch einheitenlos.

### 3.9.2 Daten des Problems

Zur Beschreibung des Algorithmus der Korrespondenzbildung muss zunächst das Format des Kraterkataloges definiert werden:

**Definition 3.12 (Kraterkatalog)** Es sei  $\mathbf{R}^{[\kappa]}$  ein Satz von  $N_p$  konfigurierten kamerafesten Kraterdaten,  $\bar{\rho}$  die Einheitsradien der beobachteten Krater und  $\lambda$  die aus der Oberflächenrekonstruktion gewonnenen unskalierten Tiefenwerte entlang der Sichtstrahlen zu den beobachteten Kratern.

a) **Globaler Katalog** Der globale Kraterkatalog  $\mathbf{K}$  mit  $N_k > 0$  Einträgen hat die Struktur

$$\mathbf{K} := \begin{pmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \\ \vdots \\ \mathbf{k}_{N_k} \end{pmatrix} = \begin{pmatrix} x_{1,w} & y_{1,w} & z_{1,w} & \rho_1 \\ x_{2,w} & y_{2,w} & z_{2,w} & \rho_2 \\ & & \vdots & \\ x_{N_k,w} & y_{N_k,w} & z_{N_k,w} & \rho_{N_k} \end{pmatrix}.$$

Die Koordinaten  $(x_{1,w}, y_{1,w}, z_{1,w})$  sind die kartesischen Koordinaten des weltfesten Mittelpunktes des Kraterrandes,  $\rho$  ist sein weltfester (tatsächlicher, metrischer) Radius.  $\mathbf{K}$  ist konstant und gegeben.

b) **Lokaler Katalog** Die Beobachtungsdaten  $\bar{\rho}$ ,  $\mathbf{R}^{[\kappa]}$  und  $\lambda$  werden abgebildet auf kamerafeste Punkte und definieren einen Katalog  $\bar{\mathbf{K}}$  der beobachteten Krater:

$$\bar{\mathbf{K}} := \begin{pmatrix} \bar{\mathbf{k}}_1 \\ \bar{\mathbf{k}}_2 \\ \vdots \\ \bar{\mathbf{k}}_{N_p} \end{pmatrix} = \begin{pmatrix} \lambda_1 \hat{c}_{1,1} & \lambda_1 \hat{c}_{1,2} & \lambda_1 \hat{c}_{1,3} & \lambda_1 \bar{\rho}_1 \\ \lambda_2 \hat{c}_{2,1} & \lambda_2 \hat{c}_{2,2} & \lambda_2 \hat{c}_{2,3} & \lambda_2 \bar{\rho}_2 \\ & & \vdots & \\ \lambda_{N_p} \hat{c}_{N_p,1} & \lambda_{N_p} \hat{c}_{N_p,2} & \lambda_{N_p} \hat{c}_{N_p,3} & \lambda_{N_p} \bar{\rho}_{N_p} \end{pmatrix}$$

Dabei sind  $\hat{c}_{i,j}$  die jeweils  $j$ -ten Komponenten des  $i$ -ten konfigurierten Zentrenvektors aus  $\mathbf{R}^{[\kappa]}$ .

Diese Definition liefert zwei Listen von Zeilenvektoren (also  $N_p \times 4$ - bzw.  $N_k \times 4$ -Matrizen), von denen bekannt ist, dass eine (möglicherweise leere) skalierte und umgeordnete Untermenge der Zeilen der einen auf eine Untermenge der Zeilen der anderen abbildbar ist. Durch Messfehler in der Beobachtung ist diese Abbildung natürlich fehlerbehaftet.

Formal ist das zu lösende Korrespondenzproblem zwischen den so definierten Katalogen das Finden einer Zuordnungsmatrix (Projektion und Permutation)  $P$ , und einer skalierenden linearen Abbildung  $(H; T)$ , die die

### 3.9.3 Affine Invarianten der Katalogelemente

---

Zeilen von  $\overline{\mathbf{K}}$  in eine Untermenge der Zeilen von  $\mathbf{K}$  überführen:

$$H(\overline{P\mathbf{K}})^T + T = (P\mathbf{K})^T \quad (3.112)$$

Über  $P$  reduziert man  $\mathbf{K}$  auf seine zu Zeilen in  $\overline{\mathbf{K}}$  korrespondierenden Zeilen und permutiert sie in die korrekte Reihenfolge, mittels  $\overline{P}$  reduziert man  $\overline{\mathbf{K}}$  auf dieselbe Schnittmenge (permutiert aber nicht mehr). Über  $H$  skaliert man sie auf die selbe Größenordnung und rotiert die drei kartesischen Komponenten in eine kongruente Lage.  $T$  verschiebt die Punktwolken dann in einen gemeinsamen Ursprung.

Probleme der Form (3.112) treten so oder in ähnlicher Form in einer Vielzahl von Anwendungsgebieten auf und werden im Kontext der reinen Objektpunkt-Zuordnung (3D) unter dem Begriff *Point Set Registration* zusammengefasst. Im Laufe der Entwicklung des maschinellen Sehens wurden einige „Meilenstein-Verfahren“ erfunden, die weite Verbreitung vor allem in der Bildregistrierung und der robotischen Objekterkennung und -Klassifizierung fanden.

Für das Korrespondenzproblem von unterschiedlich dichten und großenteils oder vollständig disjunkten dreidimensionalen Punktwolken wurde Anfang der neunziger Jahre das heute immer noch sehr populäre *Iterative Closest Point*-Verfahren entwickelt (Besl und McKay, 1992). Dieses ist jedoch sehr anfällig gegenüber Ausreißern und ist iterativ nicht-linear, benötigt also eine sehr akkurate Startschätzung. Wenige Jahre später wurde die sogenannte *Softassign*-Methode erfunden (Rangarajan, Chui und Bookstein, 1997), und darauf aufbauend das *Robust Point Matching*-Verfahren (Gold u. a., 1998), die durch Berechnung einer doppelt-stochastischen Matrix (entspricht  $P$  in (3.112)) eine Ausreißerbehandlung einführten. Die Performanz auch dieser Verfahren basiert jedoch auf der Annahme, dass die zu registrierenden Punktmengen einen ähnlichen Umfang aufweisen<sup>13</sup> und vor allem die gesuchte „Struktur“ der einen Punktmenge bei Registrierung auf der anderen bezüglich der betrachteten Gütefunktionen ein signifikantes Maximum erzeugt.

Das in dieser Arbeit angetroffene Zuordnungsproblem erfüllt leider keines dieser Kriterien. Zum einen ist die Verteilung lunarer Krater in praktisch jedem gewählten Ausschnitt und auf allen relevanten Skalen selbstähnlich (Neukum, 1977; Neukum, König und Arkani-Hamed, 1975). Zum anderen ist die Erweiterung der oben genannten Verfahren auf die höhere Dimension der Katalogeintrag-Zuordnung nicht trivial.

---

<sup>13</sup>Gold u. a. testeten nur  $p_d, p_s < 1$  (Gold u. a., 1998, S.22), beim vorliegenden Problem tritt insbesondere der Fall  $p_s \gg 1$  auf.

### 3.9.3 Affine Invarianten der Katalogelemente

Diese obenstehenden Verfahren haben alle gemeinsam, dass sie auf dreidimensionalen Punktwolken operieren. Im vorliegenden Fall ist aber zu jedem Element der Kataloge ein weiteres Datum verfügbar: Der Radius des Kraters. Beide Kataloge sind nominell im selben Koordinatensystem verortet und über die Beziehung (3.112) verwandt. Die Abbildung  $H$  hat die Form

$$H = \begin{pmatrix} \Lambda R & 0 \\ 0 & \Lambda \end{pmatrix}, \quad (3.113)$$

wobei  $R$  eine Rotation in drei Dimensionen ist und  $\Lambda$  die unbekannte Welt-skala. Die Operation der rechten Seite in (3.112) ist also eine bijektive affine Abbildung vom  $\mathbb{R}^3 \times \mathbb{R}^+$  in sich selbst. Alle affinen Abbildungen sind insbesondere längen- und damit winkeltreu, ändern also diese Beziehungen innerhalb der einzelnen Kataloge nicht. Auch die Radienverhältnisse innerhalb eines Kataloges werden durch  $(H; T)$  nicht geändert. Alle solche Größen, die aus relativen Längen, Abständen, Winkeln oder Radienverhältnissen abgeleitet werden, sind also invariant zwischen den (korrekt permutierten) Katalogen. Auf dieser Tatsache baut das unten vorgestellte Verfahren auf.

Die folgenden Unterabschnitte erklären dieses Verfahren in Algorithmenfragmenten. Es wird zunächst die Bestimmung der Zuordnungsmatrizen  $P$  und  $\bar{P}$  behandelt. Anschaulich geht es darum, diejenigen Einträge zu identifizieren, die sowohl im beobachteten Katalog  $\bar{\mathbf{K}}$  vorliegen als auch im gegebenen Katalog  $\mathbf{K}$ , unabhängig von ihrer Sortierung. Dazu wird anders als bei den oben genannten Methoden außer der drei räumlichen Dimensionen auch die „Extradimension“ der Krater-Radieninformation genutzt. Das Gesamtverfahren zur Korrespondenzgenerierung teilt sich konzeptionell in drei Kernschritte, die auf eine Lösung von (3.112) führen:

1. *Hypothesengenerierung.* Über Vergleich affiner Invarianten von Paaren und Tripeln von Katalogeinträgen wird eine Hypothese darüber aufgestellt, welche Elemente in  $\bar{\mathbf{K}}$  zu welchen Elementen in  $\mathbf{K}$  korrespondieren. Resultat sind die hypothetische Zuordnungsmatrizen  $P$  und  $\bar{P}$ .
2. *Hypothesenüberprüfung.* Diese Hypothese wird überprüft durch Vergleich der weiteren und abgeleiteten Daten der Katalogelemente und für plausibel befunden oder auch nicht.
3. *Hypothesenerweiterung.* Ist eine plausible Hypothese gefunden, wird sie erweitert um ein weiteres Korrespondenzpaar aus den Katalogen, und wiederum überprüft. Die Matrizen  $P$  und  $\bar{P}$  werden dabei erweitert.

### 3.9.4 $k$ -vector-Ansatz zur Hypothesengenerierung

---

Dies wird solange fortgesetzt, bis ausreichend Korrespondenzen gefunden sind, um die Matrix  $H$  und die Translation  $T$  zu berechnen.

4. *Berechnung der Kamerapose.* Bei einer ausreichenden Anzahl von Korrespondenzen kann die Kamerapose berechnet werden.

Im folgenden Unterabschnitt wird zunächst mit dem Aufstellen der initialen Hypothese begonnen.

### 3.9.4 $k$ -vector-Ansatz zur Hypothesengenerierung

Ein mit der Kraterkatalog-Zuordnung verwandtes und schon gelöstes Problem ist die Identifizierung und Zuordnung von Sternen im Betrieb von Star Trackern zur hochgenauen Lagebestimmung von Raumfahrzeugen. Da Sterne immer „im Unendlichen“ abgebildet sind<sup>14</sup>, fällt in dem Falle die unbekannte Skala  $\lambda$  weg und die Transformation  $H$  ist eine reine Rotation.

Zur Lösung dieses Zuordnungsproblems konstruierten Samaan, Mortari und Junkins ein Verfahren, das die unbekannte Rotation dadurch vernachlässigt, dass es auf rotationsinvarianten Eigenschaften von *Konstellationen* von Kratern aufbaut. Insbesondere wird auf einem Teil der Potenzmenge der Punktwolken eine implizite Ordnungsrelation definiert, die einen der Freiheitsgrade des kombinatorischen Grundproblems in eine lineare Suche überführt (Samaan, Mortari und Junkins, 2006). Der selbe Ansatz wird auch in vorherigen Arbeiten zur Lösung des Krater-Korrespondenzproblems verwendet, wobei dort immer von striktem Nadir Pointing ausgegangen wird (Hale, 2007; Hanak, 2009).

Bezeichnet wird der Ansatz als „ $k$ -vector range search technique“ und basiert vereinfacht auf dem folgenden Prinzip: Für alle möglichen Dreiecke über die Menge von Punkten eines Sternkataloges wird der jeweils kleinste Innenwinkel notiert. Alle diese Dreiecke werden bezüglich dieses kleinsten Innenwinkels aufsteigend im sogenannten  $k$ -vector sortiert. Über die wesentlich kleinere Menge der in einem Bild beobachteten Sterne werden sequentiell wieder Dreiecke gebildet und ihr kleinster Innenwinkel im  $k$ -vector gesucht (mit geeigneten Fehlermargen, daher „range search“). Ist ein Kandidat-Dreieck aus Katalogsternen gefunden, das einen sehr ähnlichen kleinsten Innenwinkel zu dem Gesuchten besitzt, werden die restlichen dazugehörigen Informationen verglichen, wie relative Helligkeit der Sterne und ein weiterer Innenwinkel. Passen diese mit den beobachteten Werten zusammen, so ist eine Korrespondenz

---

<sup>14</sup>Es wird durch die üblicherweise zurückgelegten Distanzen noch keine Parallaxe oberhalb der Auflösungsfähigkeit der Sternkamera erzeugt.



von drei Sternen gefunden, aus denen sich die gesuchte Rotation errechnen lässt.

Die Suche im  $k$ -vector wird dabei erleichtert durch den Umstand, dass die kleinsten Innenwinkel offenbar robust linear verteilt sind, also durch ein einfaches lineares Modell und geeignete Suchintervalle sich die Suche von kombinatorischem auf linearen Aufwand reduzieren lässt.

Es werden während dieses Vorgangs also zwei der zum Ende des letzten Unterabschnitts genannten konzeptionellen Schritte gemacht: Zunächst wird über die Suche des kleinsten Innenwinkels im  $k$ -vector eine *Hypothese* generiert darüber, welchen Katalogsternen die drei beobachteten Sterne zugeordnet werden sollten. Dann wird über Vergleich der weiteren oben genannten unabhängigen Information zu den einzelnen Sternen diese Hypothese *überprüft* und dann bestätigt oder auch verworfen.

Es zeigt sich, dass dieses Verfahren mit leichten Modifikationen als Startpunkt für die Zuordnung des Kraterkatalogs und des Katalogs der beobachteten Krater genutzt werden kann. Bei der Kraterkatalog-Zuordnung sind keine zweidimensionalen Punktmengen Gegenstand, wie es bei Sternen wegen ihrer vernachlässigbaren Parallaxe der Fall ist, sondern dreidimensionale Punktwolken. Dreiecke lassen sich natürlich trotzdem über sie bilden und damit lässt sich auf dieselbe Weise der  $k$ -Vektor aufbauen. Die reine Hypothesengenerierung lässt sich also übernehmen. Für den Anwendungsfall sei der  $k$ -vector über den Katalog  $\mathbf{K}$  wie folgt definiert:

**Definition 3.13 ( $k$ -vector)** Zum gegebenen Kraterkatalog  $\mathbf{K}$  aus Def. 3.12 seien Winkel-Index-Elemente

$$\vec{k}_m := (\alpha_m, i_m, j_m, k_m), \quad i_m, j_m, k_m \leq N_k \text{ und paarweise ungleich,}$$

definiert und damit der  $k$ -vector der Länge  $N_{\vec{k}} \leq \binom{N_k}{3}$  als

$$\vec{k} := (\vec{k}_1, \dots, \vec{k}_{N_{\vec{k}}}).$$

Die Indizes  $i_m$ ,  $j_m$  und  $k_m$  der Elemente  $\vec{k}_m$  seien so sortiert, dass die Winkel zwischen den Strecken, die von den durch sie indizierten Punkten ausgehen,

### 3.9.4 $k$ -vector-Ansatz zur Hypothesengenerierung

aufsteigend sortiert sind (und der kleinste von ihnen  $\alpha_m$  ist):

$$\alpha_m := \begin{array}{l} \triangleleft ((\mathbf{k}_{j_m} - \mathbf{k}_{i_m})C_{\text{pos}}, (\mathbf{k}_{k_m} - \mathbf{k}_{i_m})C_{\text{pos}}) \\ \leq \triangleleft ((\mathbf{k}_{i_m} - \mathbf{k}_{j_m})C_{\text{pos}}, (\mathbf{k}_{k_m} - \mathbf{k}_{j_m})C_{\text{pos}}) \\ \leq \triangleleft ((\mathbf{k}_{i_m} - \mathbf{k}_{k_m})C_{\text{pos}}, (\mathbf{k}_{j_m} - \mathbf{k}_{k_m})C_{\text{pos}}) \end{array} \quad \left| \quad C_{\text{pos}} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right.$$

Die  $k$ -vector-Elemente  $\vec{k}_m$  seien ihrerseits aufsteigend sortiert in  $\vec{\mathbf{k}}$  nach  $\alpha_m \leq \alpha_{m+1}$  für alle  $1 \leq m < N_{\vec{\mathbf{k}}}$ .

Die tatsächliche Länge  $N_{\vec{\mathbf{k}}}$  des so definierten  $k$ -vectors hängt davon ab, ob man *alle* Tripel über  $\mathbf{K}$  zulässt oder einige von vornherein ausschließt in Abhängigkeit von Eigenschaften wie den Längenverhältnissen der Dreiecksseiten, den Radienverhältnissen der referenzierten Krater, etc. Für die Konstruktion des Algorithmus in diesem Abschnitt ist das aber nicht weiter relevant.

Liegt ein solcher  $k$ -vector zu einem Katalog  $\mathbf{K}$  vor, müssen die ersten drei Korrespondenzhypothesen zum Beobachtungskatalog  $\vec{\mathbf{K}}$  erzeugt werden. Letzterer ist erheblich kürzer als  $\mathbf{K}$ , da er immer nur die momentan beobachteten Krater enthält, bzw. diejenigen unter ihnen, für die eine Tiefeninformation über die Oberflächenrekonstruktion erzeugt werden konnte. Daher ist es möglich, alle möglichen Tripel  $\vec{\mathbf{k}}_i, \vec{\mathbf{k}}_j, \vec{\mathbf{k}}_k$  zu bilden und nach Korrespondenzen zu suchen.

Zu jedem solchen Tripel wird zunächst eine Anzahl von Kandidaten-Einträgen  $\vec{k}$  aus  $\vec{\mathbf{k}}$  gesucht, wie beschrieben im Algorithmus 6. Da der  $k$ -vector per Konstruktion aufsteigend sortiert ist bezüglich der Minimalwinkel der referenzierten Tripel, ist diese Suche linear<sup>15</sup>.

Dieser reduzierte  $k$ -vector  $\vec{\mathbf{k}}^{(ijk)}$  referenziert über seine Elemente  $\vec{k}_m^{(ijk)}$  seinerseits alle Tripel  $\mathbf{k}_{i_m}, \mathbf{k}_{j_m}, \mathbf{k}_{k_m}$  von Einträgen aus  $\mathbf{K}$ , die als Korrespondenzen für  $\vec{\mathbf{k}}_i, \vec{\mathbf{k}}_j, \vec{\mathbf{k}}_k$  infrage kommen. Bis zu diesem Punkt wurden in Übereinstimmung mit dem Referenzverfahren aus der Literatur nur die minimalen Innenwinkel der Tripel verglichen. Um die u.U. vielen möglichen Korrespondenztripel  $\vec{k}_m^{(ijk)}$  vorzusortieren, wird ein auf den Anwendungsfall zugeschnittenes Abstands- bzw. Gütemaß benötigt, da die Eigenschaften der Kraterkataloge nicht identisch mit denen der Sternkataloge sind.

<sup>15</sup>Die Suche kann sogar in beinahe konstanter Zeit durchgeführt werden, wenn man ein Modell zur Vorhersage der Position des gesuchten Winkels in  $\vec{\mathbf{k}}$  ansetzt, anstatt den  $k$ -vector komplett zu durchsuchen. In der Referenzpublikation wird dazu ein lineares Modell über den Index verwendet. In konkreten Implementationen wird dieser Schritt immer gegangen, für die vorliegende Beschreibung des Algorithmus sind die Details nicht interessant, da das Ergebnis dasselbe ist.

```

1 Function  $k\_vector\_search\_range(\vec{k}, \bar{k}_i, \bar{k}_j, \bar{k}_k)$ 
2    $\bar{\alpha} \leftarrow \min \left\{ \begin{array}{l} \angle((\bar{k}_j - \bar{k}_i)C_{pos}, (\bar{k}_k - \bar{k}_i)C_{pos}), \\ \angle((\bar{k}_k - \bar{k}_j)C_{pos}, (\bar{k}_i - \bar{k}_j)C_{pos}), \\ \angle((\bar{k}_i - \bar{k}_k)C_{pos}, (\bar{k}_j - \bar{k}_k)C_{pos}) \end{array} \right\} \quad /* C_{pos} \text{ Def. 3.13} \\
3   \quad */ \\
4   \vec{k}^{(ijk)} \leftarrow () \\
5   \textbf{forall } \vec{k}_m \textbf{ where } |\bar{\alpha} - \alpha_m| < tol \textbf{ do} \\
6   \quad \vec{k}^{(ijk)} \leftarrow \textbf{append}(\vec{k}^{(ijk)}, \vec{k}_m) \\
7   \textbf{return } \vec{k}^{(ijk)}$ 
```

**Algorithmus 6 :** Extraktion des Bereiches des  $k$ -vectors, dessen kleinster Innenwinkel nah genug an dem des getesteten Tripels von Beobachtungen liegt

Die weiteren zur Verfügung stehenden Eigenschaften der Tripel, nämlich zwei der Innenwinkel (der dritte ist die Differenz aus  $180^\circ$  und der Summe der anderen beiden) und die Verhältnisse der Radien zueinander<sup>16</sup>, sind leider zum einen nicht annähernd gleich skaliert, zum anderen sind eventuelle Fehler in ihnen korreliert durch die in beide eingehenden Fehler der Oberflächenrekonstruktion.

Um diese Einflüsse zu minimieren, sei das folgende Abstandsmaß für zwei Tripel definiert:

**Definition 3.14 (Abstandsmaß für Kratertripel)** Es sei  $\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_k$  (bzw.  $\bar{k}_i, \bar{k}_j, \bar{k}_k$ ) ein Tripel von Einträgen eines der Kataloge gemäß Def. 3.12.

a) **Charakteristische Werte eines Tripels** Die Funktion

$$f_{chr} : (\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_k) \mapsto (\alpha_{min}, \alpha_{med}, \rho_{min}/\rho_{max}, \rho_{med}/\rho_{max}),$$

die solch ein Tripel von Katalogeinträgen (globaler oder beobachteter) auf das Tupel aus kleinstem und mittlerem Innenwinkel und Verhältnis von kleinstem zur größtem und mittlerem zu größtem Kraterradius abbildet, wird die Abbildung auf die charakteristischen Werte des Tripels genannt.

b) **Charakteristische Matrix des  $k$ -vectors** Die Abbildung des  $k$ -

<sup>16</sup>Es ist notwendig, relative Größen zu verwenden, da die absolute Skala der Beobachtungen ja nicht aufgelöst werden kann durch die Oberflächenrekonstruktion.

### 3.9.4 $k$ -vector-Ansatz zur Hypothesengenerierung

vectors  $\vec{\mathbf{k}}$  über den globalen Katalog  $\mathbf{K}$  auf die charakteristischen Werte aller seiner Tripel  $\vec{k}_m = (\alpha_m, i_m, j_m, k_m)$

$$\vec{\mathbf{k}}_{\text{chr}} := \begin{pmatrix} f_{\text{chr}}(\mathbf{k}_{i_1}, \mathbf{k}_{j_1}, \mathbf{k}_{k_1}) \\ f_{\text{chr}}(\mathbf{k}_{i_2}, \mathbf{k}_{j_2}, \mathbf{k}_{k_2}) \\ \vdots \\ f_{\text{chr}}(\mathbf{k}_{i_{N_{\vec{\mathbf{k}}}}}, \mathbf{k}_{j_{N_{\vec{\mathbf{k}}}}}, \mathbf{k}_{k_{N_{\vec{\mathbf{k}}}}}) \end{pmatrix} = \begin{pmatrix} \alpha_{\min,1} & \alpha_{\text{med},1} & \frac{\rho_{\min,1}}{\rho_{\max,1}} & \frac{\rho_{\text{med},1}}{\rho_{\max,1}} \\ & & \vdots & \\ \alpha_{\min,N_{\vec{\mathbf{k}}}} & \alpha_{\text{med},N_{\vec{\mathbf{k}}}} & \frac{\rho_{\min,N_{\vec{\mathbf{k}}}}}{\rho_{\max,N_{\vec{\mathbf{k}}}}} & \frac{\rho_{\text{med},N_{\vec{\mathbf{k}}}}}{\rho_{\max,N_{\vec{\mathbf{k}}}}} \end{pmatrix}$$

beschreibt seine charakteristische Matrix

$$S_{\vec{\mathbf{k}}} := \left( \left( \vec{\mathbf{k}}_{\text{chr}} \right)^T \vec{\mathbf{k}}_{\text{chr}} \right)^{-1}.$$

#### c) Charakteristischer Abstand von Tripeln Zur Differenz

$$\mathbf{d} := f_{\text{chr}}(\mathbf{k}_{i_m}, \mathbf{k}_{j_m}, \mathbf{k}_{k_m}) - f_{\text{chr}}(\bar{\mathbf{k}}_i, \bar{\mathbf{k}}_j, \bar{\mathbf{k}}_k)$$

eines durch einen  $k$ -vector-Eintrag  $\vec{k}_m$  referenzierten Tripels aus  $\mathbf{K}$  und eines Tripels aus dem Katalog der Beobachtungen  $\bar{\mathbf{K}}$  ist die Funktion  $d_{\text{chr}}$  mit

$$d_{\text{chr}} \left( \vec{k}_m, (\bar{\mathbf{k}}_i, \bar{\mathbf{k}}_j, \bar{\mathbf{k}}_k) \right) := \sqrt{\mathbf{d}^T S_{\vec{\mathbf{k}}} \mathbf{d}}$$

der „verallgemeinerte quadratische Punktabstand“, bzw. der MAHALANOBIS-Abstand der charakteristischen Werte der Tripel.

Diese Definition des Abstandes zweier Tripel voneinander berücksichtigt die tatsächliche Verteilung und Skalierung ihrer charakteristischen Werte und eventuelle Korrelationen zwischen ihnen. Seine Verwendung ist solange wohldefiniert, wie für die Bildung von Tripeln über die Beobachtungen dieselben Ausschlusskriterien verwendet werden, die auch in die Aufstellung des  $k$ -vectors eingeflossen sind, also die verglichenen Punkte durch denselben Prozess generiert werden<sup>17</sup>. Optimal ist dieses Abstandsmaß natürlich nur für Daten, die entlang ihrer Hauptachsen bereits normalverteilt sind. Das trifft im vorliegenden Fall nicht zu. Dennoch überwiegt die Spezifität dieser Sortierung die einer reinen Betrachtung nach dem Abstand des kleinsten Innenwinkels eines Dreiecks erheblich, insbesondere, wenn Kataloge mit sehr vielen Elementen

<sup>17</sup>Als Grundgesamtheit kann man hier die 3er-Potenzmenge über *alle* tatsächlich existierenden Krater überhaupt betrachten. Der  $k$ -vector ist eine sortierte Stichprobe aus dieser.

vorliegen, die viele Dreiecke mit sehr ähnlichen Winkeln enthalten.

Algorithmus 7 nutzt diese Abstandsmetrik, um den in Algorithmus 6 gefundenen Suchbereich  $\vec{\mathbf{k}}^{(ijk)}$  der Plausibilität nach zu sortieren. Ergebnis des Algorithmus ist Vektor  $\Pi$  ausführlicher Permutationen  $\pi$ , die die Korrespondenzen über ihre Katalogindizes eindeutig festhalten. Algorithmus 7

```

1 Function k_vector_initial_hypos( $\vec{\mathbf{k}}, \bar{\mathbf{k}}_i, \bar{\mathbf{k}}_j, \bar{\mathbf{k}}_k$ )
2    $\vec{\mathbf{k}}^{(ijk)} \leftarrow \mathbf{k\_vector\_search\_range}(\vec{\mathbf{k}}, \bar{\mathbf{k}}_i, \bar{\mathbf{k}}_j, \bar{\mathbf{k}}_k)$       /* Alg. 6 */
3    $\vec{\mathbf{k}}^{(ijk)} \leftarrow \mathbf{sort\_kvec\_by\_dist\_func}(d_{chr}(\vec{\mathbf{k}}^{(ijk)}, (\bar{\mathbf{k}}_i, \bar{\mathbf{k}}_j, \bar{\mathbf{k}}_k)))$ 
4    $\{i', j', k'\} \leftarrow \mathbf{sort\_by\_angle\_at\_indexed\_node}(i, j, k, \vec{\mathbf{K}})$ 
5    $\Pi \leftarrow ()$ 
6   forall  $\vec{\mathbf{k}}_m^{(ijk)} : m = 1, \dots, \text{length}(\vec{\mathbf{k}}^{(ijk)})$  do
7      $\pi \leftarrow \begin{pmatrix} i' & j' & k' \\ i_m & j_m & k_m \end{pmatrix}$ 
8      $\Pi \leftarrow \mathbf{append}(\Pi, \pi)$ 
9 return  $\Pi$ 

```

**Algorithmus 7 :** Alle initialen Korrespondenzhypothesen als Vektor von Permutationen aufsteigend sortiert nach ihrem charakteristischen Abstand vom Beobachtungstriple

verwendet zwei nicht vorher definierte Funktionen:

**sort\_kvec\_by\_dist\_func():** Sortiert den  $k$ -vector-Bereich  $\vec{\mathbf{k}}^{(ijk)}$  aufsteigend nach dem MAHALANOBIS-Abstand seiner Werte  $\vec{\mathbf{k}}_m^{(ijk)}$  vom aktuell betrachteten Beobachtungstriple  $(\bar{\mathbf{k}}_i, \bar{\mathbf{k}}_j, \bar{\mathbf{k}}_k)$ .

**sort\_by\_angle\_at\_indexed\_node():** Sortiert das Indextripel  $(i, j, k)$  der aktuell betrachteten Elemente des Beobachtungskataloges nach dem Innenwinkel, der am durch  $i, j$  und  $k$  jeweils indizierten Punkt anliegt.

Die hier vorgestellte konkrete Methode, für ein beliebiges zulässiges Triple aus dem Katalog der Beobachtungen im oben konstruierten Sinne günstig vorsortierte korrespondierende Triple aus dem globalen Katalog zu finden, stellt den Startpunkt des Gesamtverfahrens dar. Die resultierenden Permutationen  $\pi$  haben zwei Zeilen und  $N_p$  Spalten und man kann wie folgt die Matrix  $P$  aus Gleichung (3.112) daraus konstruieren:

$$\mathbb{R}^{2 \times N_p} \ni P := \begin{cases} 1 & \pi_{1i}\text{-te Zeile, } \pi_{2i}\text{-te Spalte, } i = 1, \dots, N_p \\ 0 & \text{sonst} \end{cases} \quad (3.114)$$

Diese Starthypothesen  $\Pi$  bzw.  $P$  können durch Hinzufügen von Spalten

### 3.9.5 Bestimmung einer Kamerapose

---

erweitert werden auf weitere Korrespondenzen.  $P$  wird dann  $N_h \times N_p$  groß, die  $\pi$  wachsen auf  $2 \times N_h$ .

### 3.9.5 Bestimmung einer Kamerapose

An diesem Punkt liegt eine Reihe von Hypothesen über jeweils  $N_h = 3$  Korrespondenzen vor. Erst ab  $N_h \geq 4$  Punktkorrespondenzen lässt sich aber die Validität der Hypothesen robust überprüfen<sup>18</sup>, nämlich durch das Lösen der perspektivischen Projektionsgleichungen nach der Kamerapose und Verwenden dieser zur Rückprojektion der hypothetischen Korrespondenzen im globalen Katalog ins Bild, woraufhin die Residuen gegen die Beobachtungen berechnet und bewertet werden können.

Es ist daher nötig, ein Verfahren zu implementieren, dass das sogenannte  $PnP$  (Perspektivisches  $n$ -Punkte-Problem) löst. Das  $PnP$  beschreibt in seiner allgemeinsten Form die Lösung eines Gleichungssystems  $\alpha_k x_k = Ay_k$ , für  $x_k$  Bildpunkte und  $y_k$  Raumpunkte im Kamerakoordinatensystem. Im Kontext der allgemeinen Kameraposenbestimmung ist die klassische Methode zur Lösung dieses Problems die DLT (Direct Linear Transformation)<sup>19</sup>.

Für Spezialfälle wie kalibrierte Kameras und bestimmte Anzahlen von Gleichungen (entspricht der Anzahl der Bildpunkt-Raumpunkt-Korrespondenzen) existieren aber auch eine Reihe ausgereifter Algorithmen, die nicht an den Skalierungsproblemen der DLT leiden. Das minimale  $PnP$ -Problem, „P3P“, ist effizient algebraisch lösbar, hat jedoch oft mehrere Lösungen (Gao u. a., 2003). Die Erweiterung auf vier oder mehr Korrespondenzen macht die Posenberechnung zu einem überbestimmten Problem, wodurch die Wirkung von Datenfehlern abgemildert werden kann. Andererseits steigt der Berechnungsaufwand mit der wachsenden Anzahl von Punktkorrespondenzen, es existieren Algorithmen mit Komplexitäten von  $\mathcal{O}(n^8)$  bis hinab zu  $\mathcal{O}(n)$  (Penate-Sanchez, Andrade-Cetto und Moreno-Noguer, 2013).

Einer solcher  $\mathcal{O}(n)$ -Algorithmus ist der sogenannte EPnP (Efficient Perspective  $n$ -Point Solver) (Lepetit, Moreno-Noguer und Fua, 2009). Er ist

---

<sup>18</sup>Rein geometrisch lässt sich für nicht-orthografische Kameras in vielen Fällen für drei nicht-symmetrisch angeordnete Punkt-Bildpunkt-Paare die Kamerapose errechnen. Für die meisten praktischen Anwendungen wird aufgrund solcher immer wieder auftretender Symmetrien und der Abbildungsfehler mindestens eine weitere Punktkorrespondenz benötigt, um die errechnete Kamerapose zu verifizieren. Bei signifikanten Fehlern in den Daten eventuell sogar mehr. Für einen guten Überblick über die Lösbarkeit solcher Probleme, vgl. (Gao u. a., 2003).

<sup>19</sup>Für eine Herleitung, vgl. (Hartley und Zisserman, 2003). Die DLT kann auch die Parameter der Kamerakalibration schätzen, dies ist im vorliegenden Fall nicht nötig.

eine Weiterentwicklung des Algorithmus von Ansar und Daniilidis, vlg. (Ansar und Daniilidis, 2003), der bereits von Ansar und Cheng in ihrer Studie zur Kraternavigation im Mars-Szenario verwendet wurde (Ansar und Cheng, 2005).

In seiner Grundform ist  $EPnP$  geeignet zur Kameraposenbestimmung für vier oder mehr Korrespondenzen, kann den planaren Fall bewältigen und generell mit schlecht skalierten Daten umgehen. Die Attraktivität des Algorithmus ist in diesen Eigenschaften begründet, dass er unabhängig von der Anzahl der Korrespondenzpunkte und der Datenskalisierung in den allermeisten Fällen eine robuste Grundlösung liefert. Er ist beispielsweise in OpenCV zur Bestimmung von Initialposen vor einer Nachiteration der DLT-Gleichungen implementiert (opencv.org, nodate). Für diese Arbeit wurde eine eigene Implementierung erstellt.

Streng genommen ist der Kern des  $EPnP$ -Algorithmus nur die Bestimmung eines Satzes von Skalaren einer Linearkombination des Nullraumes einer Systemmatrix, die die baryzentrischen Basisvektoren der beobachteten Punkte beschreiben. Die Berechnung der Pose selbst ist in der Referenzimplementierung von Lepetit, Moreno-Noguer und Fua durch Zerlegung einer Kovarianzmatrix der zu vergleichenden Punktwolken per SVD (Singular Value Decomposition) gelöst, und Ergebnis ist eine Rotationsmatrix. Im Kontext eines Navigationsverfahrens ist es aber sinnvoll, direkt mit Lagequaternionen zu arbeiten. Das für diese Arbeit gewählte Verfahren, diese direkt aus den (mittels  $EPnP$  rekonstruierten) Punktwolken zu gewinnen, ist das sogenannte QCP (Quaternion-based Characteristic Polynomial) (Theobald, 2005).

Die Funktion `camera_pose_from_correspondences()` in Alg. 8 referenziert die  $EPnP$ - und QCP-Algorithmen mittels der Funktion `EPnP_QCP_solve()`. Sie sollen hier nicht im Detail reproduziert werden, man lese diese Details in der angegebenen Literatur nach. Es werden die 2-Vektoren der Bildrahmenkoordinaten der Elemente des Beobachtungskataloges in der  $2 \times N_h$ -Matrix  $\mathbf{X}$  und die korrespondierenden 3-Vektoren der Weltkoordinaten des globalen Kataloges in der  $3 \times N_h$ -Matrix  $\mathbf{Y}$  zusammengefasst und mittels der Algorithmen gelöst.

Ergebnis ist ein Lagequaternion<sup>20</sup>  $q_w^C$ , das die Transformation des Welt- in das Kamerakoordinatensystem beschreibt und eine Position  $r_w$  der Kamera im Weltkoordinatensystem. Die Güte der resultierenden Kamerapose ( $q_w^C, r_w$ ) lässt sich über die im Alg. 8 berechneten *Reprojektionsresiduen* bewerten. Diese geben an, wie weit im Bild die mittels der errechneten Pose ins Bild

---

<sup>20</sup>Nach der Konvention gemäß (Wertz, 1978, S. 758 f.).

### 3.9.6 Korrespondenzhypothesen für erweiterte Graphen

```

1 Function camera_pose_from_correspondences( $\pi, \mathbf{K}, \overline{\mathbf{K}}, N_h$ )
2    $\mathbf{X} \leftarrow \begin{pmatrix} \hat{c}_{\pi_{11}1}/\hat{c}_{\pi_{11}3} & \dots & \hat{c}_{\pi_{1N_h}1}/\hat{c}_{\pi_{1N_h}3} \\ \hat{c}_{\pi_{11}2}/\hat{c}_{\pi_{11}3} & & \hat{c}_{\pi_{1H_n}2}/\hat{c}_{\pi_{1N_h}3} \end{pmatrix}$  /* Def. 3.12b */
3    $\mathbf{Y} \leftarrow \begin{pmatrix} (\mathbf{k}_{\pi_{21}} C_{\text{pos}})^T & \dots & (\mathbf{k}_{\pi_{2N_h}} C_{\text{pos}})^T \end{pmatrix}$  /*  $C_{\text{pos}}$  aus Def. 3.13 */
4    $(q_W^C, r_W) \leftarrow \text{EPnP\_QCP\_solve}(\mathbf{X}, \mathbf{Y})$ 
5   for  $i = 1, \dots, N_h$  do
6      $\mathbf{y}_C \leftarrow q_W^C \begin{pmatrix} \mathbf{k}_{\pi_{2i}} C_{\text{pos}} - r_W \\ 0 \end{pmatrix} \overline{q}_W^C$  /* Quaternionenrotation */
7     if  $\left\| \begin{pmatrix} \hat{c}_{\pi_{1i}1}/\hat{c}_{\pi_{1i}3} \\ \hat{c}_{\pi_{1i}2}/\hat{c}_{\pi_{1i}3} \end{pmatrix} - \begin{pmatrix} \mathbf{y}_{C,1}/\mathbf{y}_{C,3} \\ \mathbf{y}_{C,2}/\mathbf{y}_{C,3} \end{pmatrix} \right\|_2 > tol_{\text{repres}}$  /* Residuum */
8     then
9       return invalid pose
10  return  $(q_W^C, r_W)$ 

```

**Algorithmus 8 :** Berechnung einer Kamerapose aus einem Satz von Bildpunkt-Raumpunkt-Paaren

projizierten globalen Katalogpunkte von denen des Beobachtungskataloges entfernt sind (üblicherweise in Bildpixeln angegeben). Ist ein kritischer Wert nicht überschritten, terminiert der Algorithmus mit der Pose, ansonsten ohne Ergebnis.

### 3.9.6 Korrespondenzhypothesen für erweiterte Graphen

Wie im vorangegangenen Unterabschnitt erläutert, ist die Erweiterung der initialen Dreier-Hypothesen auf vier oder mehr Korrespondenzen notwendig, da sich leider zeigt, dass die durch die Bilderkennung und Rekonstruktion der Oberfläche entstehenden Fehler in den Beobachtungsdaten unter Umständen signifikant sein können<sup>21</sup> und daher die Berechnung der Kamerapose u.U. instabil wird.

Die zur Lösung dieses Problems entwickelte Erweiterung des  $k$ -vector-Ansatzes aus Unterabschnitt 3.9.4 beginnt mit jeweils einer initialen Hypothese der Korrespondenz von drei beobachteten Kratern zu drei Kratern im Katalog, die aus dem durch den Algorithmus 7 gewonnenen Suchraum gewählt wird. Zur

<sup>21</sup>Vgl. Kapitel 4, Abschnitt 4.3



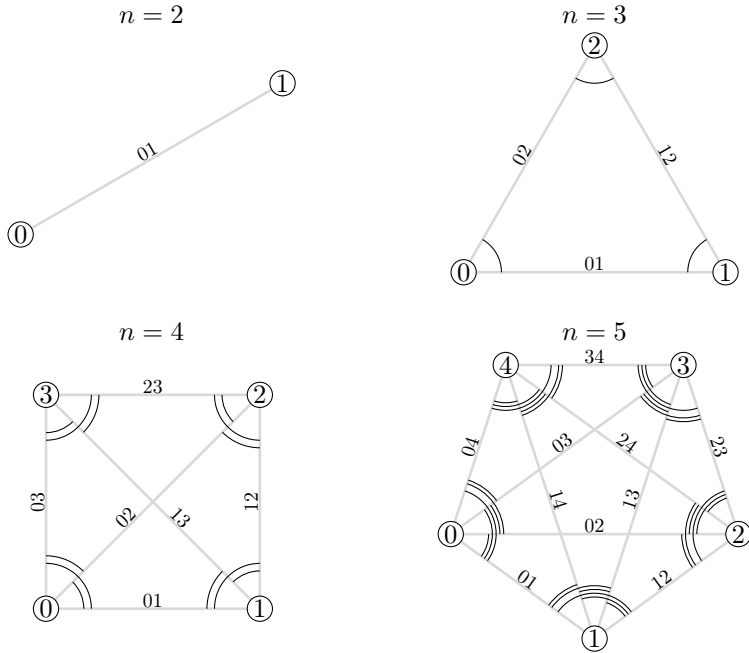


Abbildung 3.22: Abzählung innerer Kanten von kleinen Graphen und aller Winkel zwischen ihnen

robusten Schätzung einer Kamerapose wird dann eine weitere Korrespondenz benötigt. Eine solche Konstellation aus vier Kratern besitzt gegenüber dem Startdreieck mit ihren drei Innenwinkeln bereits zwölf Innenwinkel. Während man also den Kraterkatalog noch bezüglich der kleinsten Innenwinkel von Dreiecken über seine  $n = 3$  Punkte vorsortieren konnte, ist dies bei größeren Konstellationen von Punkten einerseits vom benötigten Speicher her nicht mehr praktikabel und andererseits verkleinert sich aufgrund der größeren Anzahl der Winkel und ihrer damit verbundenen Ähnlichkeit erheblich die Toleranz für Fehler. Konstellationen für  $n = 2$  bis  $n = 5$  Punkte mit ihren verbindenden Strecken und den Innenwinkeln zwischen ihnen sind in Abb. 3.22 gezeigt.

Mit steigender Anzahl  $n$  von Punkten explodiert die Anzahl der Kanten der sie verbindenden Graphen, und damit die zu vergleichenden Innenwinkel. Die Anzahl von Kanten ist  $\binom{n}{2}$ . Zur Abzählung der Kanten mit der Notation aus

### 3.9.6 Korrespondenzhypothesen für erweiterte Graphen

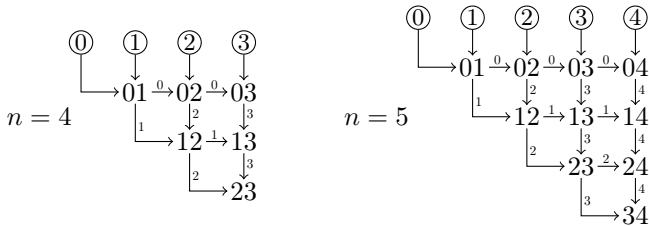
Abb. 3.22 wurde folgende Systematik konstruiert, zunächst für vier Punkte:

$n = 4$	#	0	1	2	3	4	5
	Kante	01	02	03	12	13	23

Hier ist „#“ ein Abzählungsindex. Analog ist für den in Abb. 3.22 abgebildeten Graphen von fünf Punkten:

$n = 5$	#	0	1	2	3	4	5	6	7	8	9
	Kante	01	02	03	04	12	13	14	23	24	34

Die Innenwinkel, die diese so indizierten Kanten miteinander einschließen, lassen sich über die Bildung von „Abzählungsketten“ abzählen:



Die Ketten-Indizes in den Kreisen sind gleichzeitig die Indizes der Punkte, an denen alle Kanten enden, zwischen denen Winkel gebildet werden müssen. Die Elemente der Kette sind die Bezeichner dieser Kanten. Am Fall  $n = 5$  sind diese, mit den Kantenbezeichnern und ihren Abzählungs-Indizes aus den Tabellen oben:

Kettenidx.	Kette	indiziert $(k)_{ij}$
0	01 – 02 – 03 – 04	0 1 2 3
1	01 – 12 – 13 – 14	0 4 5 6
2	02 – 12 – 23 – 24	1 4 7 8
3	03 – 13 – 23 – 34	2 5 7 9
4	04 – 14 – 24 – 34	3 6 8 9

Der  $j$ -te Kantenindex der  $i$ -ten Kette,  $k_{ij}$ , lässt sich wie folgt direkt angeben:

$$k_{ij} = \begin{cases} j = 0, \dots, i-1 & : j(n-1) + (i-1) - (j^2 + j)/2 \\ j = i, \dots, n-2 & : i(n-1) + j - (i^2 + i)/2 \end{cases} \quad (3.115)$$

Es kommen also für  $n$  Punkte mit  $\binom{n}{2}$  Kanten  $n$  solche Ketten der Länge  $n-1$  zusammen. An jedem Punkt sind also  $\binom{n-1}{2}$  Winkel zu bilden, und für

den gesamten Graphen gibt es daher  $n \binom{n-1}{2}$  Innenwinkel (Durchzählen an den Diagrammen in Abb. 3.22 bestätigt: Der Graph für den Fall  $n = 3$  hat drei Innenwinkel,  $n = 4$  hat zwölf,  $n = 5$  hat 30). Diese Abzählungsketten sind mit Hinblick auf effiziente Implementierungen absichtlich so konstruiert, dass hinzufügen eines  $(n + 1)$ -ten Punktes zu einem Korrespondenzgraphen von bisher  $n$  Punkten nur  $n$  weitere Verbindungsstrecken erzeugt, und damit  $n(n - 1) + n$  neue Winkel.

Ausgehend von den in Algorithmus 7 generierten initialen Hypothesen  $\pi$  zu je drei Korrespondenzen ist jetzt eine Erweiterung auf einen vierten, fünften oder weitere Punkte über die oben vorgestellten Abzählsystematiken möglich. Dies ist ausgeführt in Algorithmus 10. Der Parameter  $N_h$  legt fest, bei welcher Anzahl von Korrespondenzen die Hypothesenerweiterung abgebrochen wird. Bei jeder Erweiterung einer Hypothese muss eine Reihe von Plausibilitätskriterien erfüllt sein. Diese Tests sind in Algorithmus 9 zusammengefasst und werden in Algorithmus 10 verwendet. Die in Algorithmus 9 verwendeten

```

1 Function hypo_is_valid( $\pi, \mathbf{K}, \overline{\mathbf{K}}, N_h$ )
2   if not hypo_rel_radii_are_consistent( $\pi, \mathbf{K}, \overline{\mathbf{K}}, N_h$ ) then
3     return false
4   if not hypo_graph_angles_are_consistent( $\pi, \mathbf{K}, \overline{\mathbf{K}}, N_h$ ) then
5     return false
6   return true

```

**Algorithmus 9 :** Plausibilitätstests für Korrespondenzhypothesen sortiert nach aufsteigendem Rechenaufwand

Testfunktionen sind:

**hypo\_rel\_radii\_are\_consistent():** Überprüfe für beide Graphen die relativen Fehler der Verhältnisse der bisherigen Radien zum neu hinzugefügten und vergleiche das Maximum dieser Fehler mit einer Toleranzschwelle:

$$tol_{\text{relrad}} \stackrel{?}{>} \max_{i=1, \dots, N_h} \left\{ \frac{\left| \lambda_{\pi_{1i}} \bar{\rho}_{\pi_{1i}} / \lambda_{\pi_{1(N_h+1)}} \bar{\rho}_{\pi_{1(N_h+1)}} - \rho_{\pi_{2i}} / \rho_{\pi_{2(N_h+1)}} \right|}{\rho_{\pi_{2i}} / \rho_{\pi_{2(N_h+1)}}} \right\} \quad (3.116)$$

Dabei sind  $\rho_i = \mathbf{k}_i(0, 0, 0, 1)^T$  die Radiuskomponenten der  $i$ -ten Katalogzeile.

**hypo\_graph\_angles\_are\_consistent():** Berechne für beide Graphen die Innenwinkel aller Strecken (mittels der oben vorgestellten Abzählungen) und

### 3.9.7 Gesamtalgorithmus

nehme den Durchschnitt der absoluten Fehler der zwischen den Graphen korrespondierenden Winkel. Vergleiche diesen Durchschnitt gegen eine Toleranzschwelle  $tol_{\text{relang}}$ .

```
1 Function k_vector_extend_hypo( $\pi, \mathbf{K}, \overline{\mathbf{K}}, N_h, N_{h,\min}$ )
2   forall  $m = 1, \dots, N_k : m \neq \pi_{i1}, i = 1, \dots, N_h$  do
3     forall  $n = 1, \dots, N_p : n \neq \pi_{i2}, i = 1, \dots, N_h$  do
4        $\pi' \leftarrow \left( \pi \mid \begin{smallmatrix} m \\ n \end{smallmatrix} \right)$ 
5       if hypo_is_valid( $\pi', \mathbf{K}, \overline{\mathbf{K}}, N_h$ ) /* Alg. 9 */
6         then
7            $N_h \leftarrow N_h + 1$ 
8           if  $N_h \geq N_{h,\min}$  then
9              $(q_w^C, r_w) \leftarrow \text{camera\_pose\_from\_corresponden-}$ 
10               $\text{ces}(\pi', \mathbf{K}, \overline{\mathbf{K}}, N_h)$  /* Alg. 8
11              */
12              if  $(q_w^C, r_w)$  is not invalid pose then
13                return  $\pi'$ 
14              else
15                continue
16              else
17                 $\pi' \leftarrow \text{k\_vector\_extend\_hypo}(\pi', \mathbf{K}, \overline{\mathbf{K}}, N_h, N_{h,\min})$ 
18                /* Rekursion */
19      return invalid hypothesis
```

**Algorithmus 10 :** Rekursive Erweiterung einer Hypothese zwischen den Beobachtungsdaten und dem globalen Katalog mit Plausibilitätstests

Algorithmus 10 ist rekursiv und terminiert dann entweder mit einer validen Hypothese der vorher festgelegten Mindestgröße  $N_{h,\min} \geq 4$  oder mit einer Nachricht, dass keine valide Hypothese gefunden werden kann.

### 3.9.7 Gesamtalgorithmus

Mit der Vorstellung des Mechanismus zur Erweiterung der partiellen Korrespondenzhypothesen im vorangegangenen Unterabschnitt ist das Gerüst zur Beschreibung der Korrespondenzbildung zwischen dem Katalog der Beobach-

tungen und dem globalen Kraterkatalog komplett. Algorithmus 11 fügt jetzt alle bisher beschriebenen Algorithmenfragmente zusammen und beschreibt das Gesamtverfahren im Sinne dieses Abschnittes.

Eingaben des Algorithmus sind der globale Kraterkatalog, der auf dem globalen Kraterkatalog basierende  $k$ -vector  $\vec{k}$  (diese ändern sich nach ihrer Initialisierung für ein Szenario oder eine Mission nicht mehr online), der Katalog der Beobachtungen  $\overline{K}$  und der Parameter  $N_{h,min}$  der erforderlichen Anzahl von Korrespondenzen. Die resultierende Permutation  $\pi$  ist das Ergebnis des

```

1 Function catalog_correspondence_search( $\vec{k}, \mathbf{K}, \overline{K}, N_{h,min}$ )
2   forall admissible  $\overline{k}_i, \overline{k}_j, \overline{k}_k$  do
3      $\Pi \leftarrow \text{k\_vector\_initial\_hypos}(\vec{k}, \overline{k}_i, \overline{k}_j, \overline{k}_k)$       /* Alg. 7 */
4     forall  $\pi \in \{\Pi_i \mid i = 1, \dots, \text{length}(\Pi)\}$  do
5        $N_h \leftarrow 3$ 
6        $\pi \leftarrow \text{k\_vector\_extend\_hypo}(\pi, \mathbf{K}, \overline{K}, N_h, N_{h,min})$  /* Alg.
7         10 */
8       if  $\pi$  is not invalid hypothesis then
9         return  $\pi$ 
10      else
11      return no correspondences found
12 return no correspondences found

```

**Algorithmus 11 :** Suche nach Korrespondenzen für den Katalog der Beobachtungen im globalen Katalog

in dieser Arbeit vorgestellten Gesamtverfahrens. Terminiert der Algorithmus 11 mit einem gültigen Ergebnis, stellt dieses aufgrund der umfangreichen Plausibilitätsprüfungen durch die Hilfsfunktionen in Algorithmus 9 und die Prüfung der Rückprojektions-Residuen in der Kameraposen-Berechnung fast sicher eine korrekte Zuordnung dar.

Verbleibende Restunsicherheit lässt sich an diesem Punkt dadurch ausräumen, dass man die durch die Korrespondenzhypothese gewonnene Pose-schätzung nutzt, um den gesamten sichtbaren Teil des Kataloges  $\mathbf{K}$  in das Bild zu projizieren und mit *allen* erkannten Kratern zu vergleichen. Bewertung aller Residuen von durch Nachbarschaft im Bild gebildeten Paaren von Erkennungen und Rückprojektionen macht es unwahrscheinlich, eine falsche Kamerapose für richtig zu halten.

## 3.10 Zwischenfazit

Hiermit ist der theoretische Teil der vorliegenden Arbeit abgeschlossen. In der Kapiteleinleitung auf Seite 41 wurde das Ziel des zu konstruierenden Verfahrens formuliert und die Abschnitte zwischen dort und hier lieferten Schritt für Schritt die notwendigen Werkzeuge zur Erreichung dieses Ziels.

Bis jetzt bleibt das konstruierte Verfahren aber abstrakt: Was fehlt, ist eine Implementierung. Obwohl die Konstruktion der Interpolationssysteme durch Def. 3.11 auf Seite 94 eindeutig ist und die im Abschnitt 3.8 zur Konfiguration und im Abschnitt 3.9 zur Korrespondenzbildung konstruierten Algorithmen sehr detailliert sind, bleibt an vielen Stellen erhebliche Freiheit in der Wahl der numerischen Mittel zu ihrer tatsächlichen Realisierung.

Auch arbeitet das vorgestellte Verfahren an vielen Stellen approximativ, etwa in der Modellierung der Kraterränder als Kreise, der Modellierung der beobachteten Oberfläche über kubische Splines etc., mit den daraus resultierenden Approximationsfehlern. Schließlich muss man in Betracht ziehen, dass die Eingabedaten in das Verfahren nicht den Vorstellungen entsprechen. Inwiefern beeinflussen diese genannten Unwägbarkeiten die korrekte Funktion der implementierten Methoden?

Diese Fragen sind Gegenstand des folgenden Kapitels. Zu ihrer Beantwortung wurde eine Referenzimplementierung der bisher vorgestellten Algorithmen erstellt und instrumentiert, um an allen relevanten Stellen die korrekte Funktion der Teilalgorithmen verifizieren zu können. Dies und die noch fehlende repräsentative Simulation von Szenarien und ein Konzept zur Darstellung der Ergebnisse und Fehler wird auf den folgenden Seiten erörtert.

# Kapitel 4

## Fehleranalyse des Identifikationsverfahrens

### 4.1 Einleitung

Das in Kapitel 3 konstruierte Verfahren zur Rekonstruktion einer sichtbaren Oberfläche aus Messungen von Kraterzentren-Richtungen und -Normalen ist in der Praxis Fehlern der Eingangsdaten, nämlich dem Bild, und akkumulierten Fehlern der vorangehenden Berechnungsschritte, ausgesetzt. Die generierte Lösung ist daher unvermeidlich mit gewissen Ausgangsfehlern behaftet. Um die Qualität der Lösung zu beurteilen, muss das Fehlerverhalten des Zusammenspiels aller verwendeten Algorithmen verstanden sein. Zur Motivation und Einordnung dieser Fehleranalyse in den Gesamtkontext der Arbeit dient das Flussdiagramm in Abbildung 4.1 und die folgende Erläuterung der Fehlerquellen und ihrer Fortpflanzungskette.

Die in der Abbildung dargestellte Gesamtverarbeitungskette gliedert sich in mehrere Teile, deren Abgrenzung gegeneinander sich durch Transformationen zwischen, bzw. Abbildungen von verschiedenen Räumen aufeinander, begründet. Diese Abgrenzungen, bzw. Verfahrensschritte, sind:

1. Es wird aus dem Eingangsbild  $I$  ein Satz Ellipsen  $\mathbf{E}$  extrahiert, der die Ränder der sichtbaren detektierten Krater repräsentiert (Kapitel 2).
2. Aus diesen wird mithilfe des Lochkameramodells die Menge von Zentren- und Normalenvektoren  $\mathbf{R} = (\dots, ((c_i, n_i), (c'_i, n'_i)), \dots)$  rekonstruiert (Abschnitt 3.3).

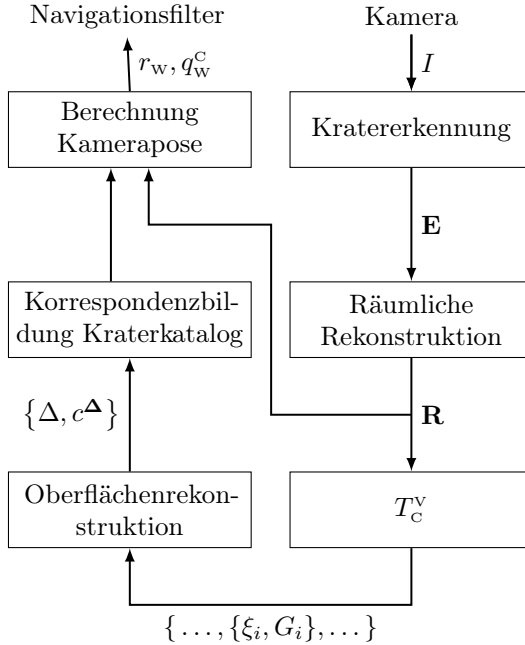


Abbildung 4.1: Datenfluss im Navigationskreis

3. Diese Menge von Vektoren werden in ihre durch die virtuellen Bildkoordinaten parametrisierte Form  $\xi_i, G_i$  transformiert (Definition 3.4 und Gleichung (3.78)).
4. Die Oberflächenrekonstruktion erzeugt daraus eine Approximationsfläche, parametrisiert durch die Triangulation und die darüber definierten Kontrollnetz-Ordinaten  $c^\Delta$  (Kapitel 3, Abschnitt 3.7).
5. Auf Basis der Approximationsfläche wird eine Korrespondenz zwischen den erkannten Kratern und einem Kraterkatalog erstellt.
6. Mit diesen Korrespondenzen zusammen bilden die Daten ein perspektivisches  $n$ -Punkte-Problem, dessen Lösung die Kamerapose ist.

In den folgenden Unterabschnitten sollen diese Einzelschritte und die in ihnen auftretenden Fehlertypen hinsichtlich ihrer zu erwartenden Größenordnungen identifiziert und modelliert werden, immer in Abhängigkeit von ebenfalls zu modellierenden „Messfehlern“ des Kratererkennungsalgorithmus.



## 4.2 Modellierung der Verfahrensfehler

Die Kette der Datenpropagation im Gesamtverfahren zur Generierung einer Kamerapose aus einem Kamerabild war in Abb. 4.1 dargestellt. Modelliert man die Fehler der Zwischenstufen des Verfahrens als additiv und beobachtet ihre Fortpflanzung, dann lässt sich diese Propagationskette in zwei Unterbereiche teilen: In 1) einen Bereich der strikten Propagation reeller Größen in den Zwischenergebnissen, und 2) einen Bereich kombinatorischer Propagation.

Letzteres bedeutet in diesem Zusammenhang, dass Änderungen in den Eingangsfehlern unstetig wirken. Etwa, indem sie die Triangulierung beeinflussen und dadurch die Spline-Systeme permutieren oder sogar ihre Größe ändern. Abb. 4.2 zeigt als Erweiterung von Abb. 4.1 die Unterteilung in diese beiden Fehlerkategorien, sowie die additiven Fehlergrößen des Unterbereichs der reellen Fehlerpropagation.

Die in der Abbildung eingetragenen additiven Größen sind jeweils wieder modelliert als Summe von systematischen Fehlern der Verfahren (Index „sys“), numerischen (bzw. approximativen) Fehlern (Index „num“), oder Fehlern in den Eingangsdaten (Index „dta“). Die Fehlerart „dta“ ist genau der Messfehler im Bild und wirkt daher auf keinen der Berechnungsblöcke. Die unterscheidenden Merkmale der beiden ersten Arten von Fehlern, „sys“ und „num“, sind erstens ihre Ursachen, und zweitens das Maß, in dem sie sich durch Wahl externer Parameter beeinflussen lassen. Eine Liste möglicher Ursachen der jeweiligen Fehlerkategorien ist in Tabelle 4.1 gegeben.

Typ	Ursachen
„num“	<ul style="list-style-type: none"> <li>- Numerische Fehler der Gleichungs(-System)-Löser, bzw. gewählte Präzision</li> <li>- Gewählte Abbruchschranken iterativer Verfahren</li> </ul>
„sys“	<ul style="list-style-type: none"> <li>- Approximationsgüte der gewählten Modelle (z.B. Grad der Splinepolynome)</li> </ul>
„dta“	<ul style="list-style-type: none"> <li>- Mangelhafte Kamerakalibration</li> <li>- Unterbelichtete Bildbereiche, Sättigung, Bildrauschen</li> <li>- Sensorfehler</li> </ul>

Tabelle 4.1: Fehlerarten und -Ursachen

In Abbildung 4.2 sind einige der Fehlerquellen ausgegraut, und zwar genau dann, wenn die jeweiligen Fehlerquellen vernachlässigbar sind. „Vernachlässig-

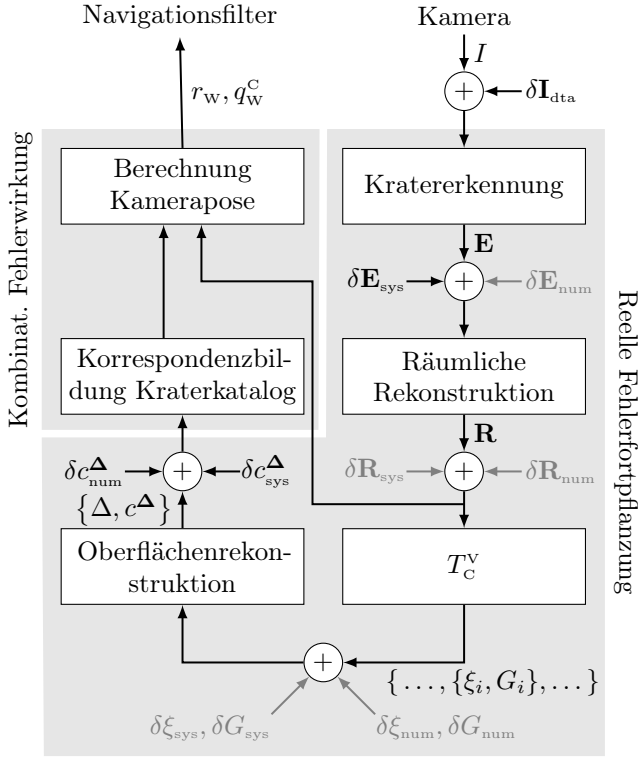


Abbildung 4.2: Fehlerquellen im Navigationskreis

bar“ ist in diesem Zusammenhang ein Teilfehler, wenn er um Größenordnungen kleiner ist als die dominanten Fehler im Gesamtverfahren. Dies trifft jeweils aus den folgenden Gründen zu:

$\delta \mathbf{E}_{\text{num}}$ : Alle in der Ellipsenapproximation verwendeten numerischen Verfahren sind direkt und deterministisch, es entsteht weder Iterationsfehler noch stochastischer Fehler. Die Ergebnissenauigkeit ist nur beschränkt durch die Rechengenauigkeit.

$\delta \mathbf{R}_{\text{num}, \text{sys}}$ : Die durch Kratererkennung generierten Ellipsen bilden direkt auf die Zentren- und Normalenvektoren ab, und umgekehrt. Es entsteht also kein Modellfehler. Die dazu zu lösenden konischen Gleichungen haben vollen

Rang, alle weiteren Ausdrücke zur Berechnung sind geschlossen lösbar. Der einzige beschränkende Faktor ist also wieder die numerische Genauigkeit, diese ist bei geeigneter Wahl der Rechengenauigkeit vernachlässigbar im oben beschriebenen Sinn.

$\delta\xi_{\text{num},\text{sys}}$  und  $\delta G_{\text{num},\text{sys}}$ : Die Kraterrand-Parameter bilden wiederum direkt (projektiv) auf die virtuellen Bildkoordinaten ab, alle Ausdrücke sind geschlossen lösbar. Es entsteht nur ein wiederum vernachlässigbarer numerischer Fehler.

In der gesamten Berechnungskette bleiben also nur noch die systematischen Fehler der Kratererkennung und die systematischen und approximativen Fehler der Oberflächenrekonstruktion relevant, mit den jeweils folgenden Ursachen:

$\delta I_{\text{data}}$ : Kein Fehler der Verarbeitungskette, sondern Messfehler in Form von Verzerrung, Sättigung, etc., als Folge von Misskalibration der Kamera, mangelhafte Parametrierung bezüglich Belichtungszeit, Strahlungsschäden am Sensor, usw..

$\delta \mathbf{E}_{\text{sys}}$ : Fehler, die durch Approximation nicht-elliptischer Kraterränder im Bild durch Ellipsen entstehen, oder auch durch Überschreiten des Bruchpunktes der Locking-Stufe des Approximationsverfahrens durch nicht zum Kraterrand gehörende Kontrastmaxima. Diese Fehler in den Ellipsen äußern sich durch Verschiebung des Ellipsenmittelpunktes im Bild, durch Stauchung/Streckung der Ellipsen-Halbachsen, und durch Rotation der Ellipse. Sie sind mittelbar als diejenigen Fehler zu verstehen, die nach räumlicher Rekonstruktion der Zentren- und Normalenvektoren aus den gestörten Ellipsen bei korrekter Kamerapose Abweichungen von den durch den Katalog definierten „wahren“ Werten erzeugen.

$\delta c_{\text{num}}^{\Delta}$ : Wird das iterative Verfahren zur Lösung der Systemgleichungen der Spline-Interpolation (Gleichungen (3.104) und (3.105)) eingesetzt, muss eine Abbruchschranke gewählt werden für die relative Änderung des Kontrollnetz-Ordinatenvektors. Die daraus resultierenden Absolutfehler des Verfahrens sind von der Größenordnung dieser Schranke.

$\delta c_{\text{sys}}^{\Delta}$ : Die Approximation der sichtbaren Oberfläche durch einen flächigen Spline hat Parameter, die seine Approximationstreue beeinflussen, nämlich hauptsächlich die Dichte der Stützstellen (Zentren- und Normalenvektoren-Paare) und den Polynomgrad des Splines. Der Grad ist fest gewählt und die Dichte der Stützstellen lässt sich nicht beeinflussen, sondern ist vom Bild abhängig.

Die oben genannten Fehler werden durch die Verarbeitungskette hindurch-

### 4.2.1 Ausgangsfehlermaße

---

propagiert, bis sie am Ausgang der Oberflächenrekonstruktion in den Abschnitt der nicht-reellen Fehlerfortpflanzung übergehen. Nach diesem Punkt sind die wirkenden Fehler nicht mehr reell, sondern diskret oder sogar binär („lösbar oder nicht“). Die Fehleranalyse in diesem Kapitel muss deswegen in zwei Schritten erfolgen: Zuerst ist die funktionelle Wirkung des Eingangsfehlers  $\delta I_{\text{dta}}$  auf die Ausgangsfehler  $\delta c^{\Delta} := \delta c_{\text{sys}}^{\Delta} + \delta c_{\text{num}}^{\Delta}$  der Splineflächen-Rekonstruktion zu untersuchen. Durch Herleitung einer Fehlerschranke oder geeignete Abdeckung des Eingangs-Fehlerraumes ergibt sich dadurch eine Verteilung der  $\delta c^{\Delta}$ . Unter ihrer Wirkung kann dann im zweiten Testteil die Erfolgswahrscheinlichkeit der Korrespondenzbildung gemessen werden.

### 4.2.1 Ausgangsfehlermaße

Als Substitut für den schlecht quantifizierbaren Bildfehler  $\delta I_{\text{dta}}$  lässt sich hier der systematische Fehler  $\delta \mathbf{E}_{\text{dta}}$  der Kratererkennung verwenden, da deren numerische Fehler wie oben besprochen vernachlässigbar sind, während ihre systematischen Fehler gerade durch die Bildfehler verursacht werden. Der entscheidende Vorteil dabei ist, dass die Ellipsen  $E_i \in \mathbf{E}$  durch nur fünf Parameter bestimmt sind. Dagegen ist das Bild  $I \in \mathcal{L}^2$  und es existiert keine offensichtlich nützliche Methode, Fehler zu messen oder für die Abdeckungstests zu generieren.

Mit Blick auf die Ausgangsfehler ist aufgrund der oben gemachten Überlegungen klar, dass nur der Fehler der Oberflächenrekonstruktion überhaupt geschlossen modellierbar ist, während der Fehler der Korrespondenzbildung aus Abschnitt 3.9 nicht-analytischer Art ist. Es wird sich jedoch im Abschnitt 4.4 herausstellen, dass Wissen über diesen Fehler ohne Kenntnis des Fehlerverhaltens der Oberflächenrekonstruktion nutzlos ist.

Die anfangs gestellte Frage nach der Wirkung von Eingangsfehlern im Bild auf die Güte der Oberflächenrekonstruktion lässt sich damit präzisieren zur Frage nach einem funktionellen Zusammenhang

$$\delta \mathbf{E}_{\text{dta}} \stackrel{?}{\mapsto} \delta c^{\Delta}, \quad (4.1)$$

bzw. nach einer Schranke für die rechte Seite unter Variation der Linken. Mit diesem Ziel vor Augen wird der Versuch gemacht, eine solche geschlossene Schranke herzuleiten.

Die oben identifizierten Eingangsfehlergrößen der Ellipsenparameter werden durch die räumliche Rekonstruktion und die Transformation in virtuelle Bildkoordinaten auf die Koeffizienten des im Abschnitt 3.7.8 konstruierten

Gleichungssystems der Oberflächenrekonstruktion abgebildet. Als Voraussetzung für die Analyse sei angenommen, dass dabei die erzeugte Triangulation nicht durch diese Fehler variiert (da ansonsten das System zusätzlich zur komponentenweisen Störung auch noch permutiert würde oder seine Größe ändern könnte). Diese Annahme ist aber nicht unrealistisch, da generell von kleinen Störungen ausgegangen wird. Die Frage nach der Art des Fehlers der Oberflächenrekonstruktion reduziert sich damit auf die Untersuchung von Eingangs- und Ausgangsstörungen der LAGRANGE-Gleichung (3.102).

Dieses System hat per Konstruktion und mit der Festsetzung des ersten Kontrollpunktes auf Null vollen Rang und daher eine eindeutige Lösung. Im Folgenden bezeichnen ganz allgemein  $A$  und  $b$  seine linke und rechte Seite, und  $x$  seine eindeutige Lösung:

$$A := \begin{pmatrix} E & e_1 & S^T & I^T \\ e_1^T & & & \\ S & & 0 & \\ I & & & \end{pmatrix}, \quad b := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ G \end{pmatrix}, \quad x := A^{-1}b. \quad (4.2)$$

Von der Lösung  $x = (c^\Delta, l_1, l_2, l_3)^\top$  ist wieder nur der erste Teil von Interesse, nicht aber die LAGRANGE-Multiplikatoren  $l_i$ . Fehler in diesen haben mit Hinblick auf das gestellte Problem keine Bedeutung, sofern sie nicht Fehler in  $c^\Delta$  bedingen. Wie man später sehen wird, ist das aber deswegen gleichzeitig auch eine Einschränkung „in Richtung der Vorsicht“, denn allenfalls *überschätzt* man den zu erwartenden Fehler.

Modelliert man die Störungen in den Koeffizienten des Problems wie oben begründet als additiv, so lässt sich die gestörte LAGRANGE-Gleichung ausdrücken als

$$(A + \delta A)(x + \delta x) = b + \delta b. \quad (4.3)$$

Der durch die Eingangsfehler  $\delta A$  und  $\delta b$  hervorgerufene Ausgangsfehler  $\delta x$  ist also nach Umformung von (4.3) mit  $Ax = b$  gerade<sup>1</sup>

$$\delta x = (A + \delta A)^{-1}(\delta b - \delta A(A^{-1}b)). \quad (4.4)$$

Welche Eigenschaft von  $\delta x$  sollte trotz der Eingangsstörungen  $\delta A$  und  $\delta b$  erfüllt sein, um im Sinne des Gesamtverfahrens als „gut“ bewertet zu werden?

<sup>1</sup>Die Annahme, dass  $A + \delta A$  nicht-singulär bleibt, ist dabei deswegen gesichert, da die Störungen  $\delta A$  nicht zufällig sind, sondern  $A + \delta A$  genau wie  $A$  selbst aus zwar gestörten aber nicht-degenerierten Eingangsgrößen konstruiert wird und  $\delta A$  die nachträgliche Differenz der jeweils wohldefinierten Lösungen zu ungestörten und gestörten Daten ist.

#### 4.2.1 Ausgangsfehlermaße

---

Man erinnere sich an den Abschnitt 3.6.2 und daran, dass die „wahre“ Oberfläche  $r(\xi)$  im Kamera-Koordinatensystem (über die virtuellen Bildkoordinaten  $\xi$ ) durch die rekonstruierte Fläche  $\hat{r}(\xi)$  selbst im optimalen Fall nur bis auf eine skalare Konstante approximiert wird:  $r(\xi) = \hat{r}(\xi)\Lambda$ . Alle nachfolgenden Methoden, die auf dieser Approximation  $\hat{r}(\xi)$  operieren, müssen sich also auf die Verhältnisse von Werten  $\hat{r}(\xi)$  untereinander beziehen, um die globale Skala  $\Lambda$  über

$$\frac{\hat{r}(\xi_i)}{\hat{r}(\xi_j)} = \frac{r(\xi_i)/\Lambda}{r(\xi_j)/\Lambda} = \frac{r(\xi_i)}{r(\xi_j)} \quad (4.5)$$

herauszukürzen und so die Referenzfläche  $r$  und die approximierte Fläche  $\hat{r}$  vergleichbar zu machen.

Die oben angesprochenen Fehler  $\delta c^\Delta$  sind genau die Fehler in den Koeffizienten des interpolierenden Splines  $s(\xi)$ , dessen Exponential  $\hat{r}$  ist, und gehen daher linear in die Auswertungen von  $s$  ein. Bezeichne  $\tilde{\hat{r}}$  die durch  $\delta c^\Delta$  gestörte Oberflächenfunktion, dann ist

$$\begin{aligned} \tilde{\hat{r}}(\xi) &= \exp(\tilde{s}(\xi)) \\ &= \exp\left(\sum_{t \in \Delta} \tilde{s}^t(\xi)\right) \\ &= \exp\left(\sum_{t \in \Delta} \sum_{\alpha} (c_{\alpha}^t + \delta c_{\alpha}^t) B_{\alpha}^t(\lambda^t(\xi))\right) \\ &= \exp\left(\sum_{t \in \Delta} \sum_{\alpha} c_{\alpha}^t B_{\alpha}^t(\lambda^t(\xi)) + \sum_{t \in \Delta} \sum_{\alpha} \delta c_{\alpha}^t B_{\alpha}^t(\lambda^t(\xi))\right) \\ &=: \exp(s(\xi) + \delta s(\xi)) \end{aligned} \quad (4.6)$$

mit einem baryzentrischen Störpolynom, das wir  $\delta s$  nennen (zur Form vgl. Def. 3.8b). Welche Art von Ausgangsstörungen  $\delta c^\Delta$  ist daher „gutartig“ in Bezug auf die betrachteten Quotienten  $\hat{r}(\xi_i)/\hat{r}(\xi_j)$ ? Insbesondere, welche ihrer Eigenschaften gewährleistet, dass

$$\left| \frac{\hat{r}(\xi_i)}{\hat{r}(\xi_j)} - \frac{\tilde{\hat{r}}(\xi_i)}{\tilde{\hat{r}}(\xi_j)} \right| \quad (4.7)$$


---

tolerierbar klein bleibt? Wegen

$$\begin{aligned}
 \frac{\tilde{r}(\xi_i)}{\tilde{r}(\xi_j)} &= \frac{\exp(\tilde{s}(\xi_i))}{\exp(\tilde{s}(\xi_j))} \\
 &= \frac{\exp(c_i + \delta c_i)}{\exp(c_j + \delta c_j)} \\
 &= \exp(c_i + \delta c_i) \cdot \exp(-(c_j + \delta c_j)) \\
 &= \exp(c_i - c_j + \delta c_i - \delta c_j) \\
 &= \frac{\exp(c_i)}{\exp(c_j)} \cdot \exp(\delta c_i - \delta c_j) \\
 &= \frac{\hat{r}(\xi_i)}{\hat{r}(\xi_j)} \cdot \exp(\delta c_i - \delta c_j)
 \end{aligned} \tag{4.8}$$

wird (4.7) zu

$$\left| \frac{\hat{r}(\xi_i)}{\hat{r}(\xi_j)} - \frac{\tilde{r}(\xi_i)}{\tilde{r}(\xi_j)} \right| = \left| \frac{\hat{r}(\xi_i)}{\hat{r}(\xi_j)} \cdot (1 - \exp(\delta c_i - \delta c_j)) \right|. \tag{4.9}$$

Bezeichne solche Quotienten von  $\hat{r}$  im Folgenden als

$$\hat{q}_{i,j} := \frac{\hat{r}(\xi_i)}{\hat{r}(\xi_j)}, \tag{4.10}$$

dann lässt sich in die typische Darstellungsform eines relativen Fehlers von  $\hat{q}_{i,j}$  Gleichung (4.9) einsetzen und wie folgt umformen:

$$\delta \hat{q}_{i,j}^{\text{rel}} := \frac{|\hat{q}_{i,j} - \tilde{q}_{i,j}|}{|\hat{q}_{i,j}|} = \left| \frac{\hat{r}(\xi_i)}{\hat{r}(\xi_j)} - \frac{\tilde{r}(\xi_i)}{\tilde{r}(\xi_j)} \right| \left| \frac{\hat{r}(\xi_j)}{\hat{r}(\xi_i)} \right| \stackrel{4.9}{=} |1 - \exp(\delta c_i - \delta c_j)|. \tag{4.11}$$

Dieser Ausdruck ist ganz offensichtlich bei  $1 - \exp(\delta c_i - \delta c_j) = 0$ , also für  $\delta c_i = \delta c_j$ , minimal.

## 4.2.2 Geschlossene Abschätzung

Resultat der obigen Überlegungen ist, dass nicht absolute normweise Größen wie  $\|\delta c^\Delta\|_\infty$  oder  $\|\delta c^\Delta\|_2$  den Ausschlag geben, sondern die Varianz der Fehler.

### 4.2.2 Geschlossene Abschätzung

Eine Abschätzung dafür im Sinne eines „Worst-Case“-Maßes ist der maximale Durchmesser von  $\delta c^\Delta$

$$\text{diam}(\delta c^\Delta) := \max_{i,j} \{|\delta c_i - \delta c_j|\} \quad (4.12)$$

und es lässt sich zusammenfassend feststellen: Wenn Fehler in den Eingangsdaten  $\delta A$  und  $\delta b$  vorliegen und dadurch eine durch  $\delta c^\Delta$  gestörte Lösung erzeugen, ist der maximale relative Fehler der Quotienten von  $\hat{r}$  beschränkt durch

$$\delta \hat{q}_{i,j}^{\text{rel}} \leq |1 - \exp(\text{diam}(\delta c^\Delta))|, \quad 1 \leq i, j \leq N_c. \quad (4.13)$$

Von Interesse ist also genau eine Abschätzung der Größe  $\text{diam}(\delta c^\Delta)$  in Abhängigkeit von  $\delta A$  und  $\delta b$ . Für die Störungstheorie linearer Systeme der Art (3.102) existiert umfangreiche Literatur. Normweise Schranken für die Lösung in Abhängigkeit von normweise beschränkten Störungen der Koeffizienten und der rechten Seite sind aber leider nicht immer informativ, insbesondere, wenn wie hier das System und die Störung selbst stark strukturiert sind. Nämlich sind per Konstruktion die Störungen  $\delta A$  und  $\delta b$  auf die folgende Art zusammengesetzt:

$$\delta A = \begin{pmatrix} \delta E & 0 & \delta S^\top & \delta I^\top \\ 0 & & & \\ \delta S & & 0 & \\ \delta I & & & \end{pmatrix}, \quad \delta b = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta G \end{pmatrix} \quad (4.14)$$

Wie man sieht, sind (blockweise betrachtet) von den  $(N_c + 1 + 3N_i + 2N_p)^2$  Elementen von  $\delta A$  schon  $(1 + 3N_i + 2N_p)^2 + 2N_c$  Elemente Null und von  $\delta b$  nur  $2N_p$  Elemente ungleich Null. Betrachtet man die Blockmatrizen intern, so bleiben noch weniger Elemente übrig:  $\delta S$  besitzt genau wie  $S$  nur  $9N_i$  von  $3N_c N_i$  Elemente ungleich Null,  $\delta I$  genau wie  $I$  nur  $6N_p$  von  $2N_c N_p$ . Die Störung ist genauso dünn besetzt wie das System selbst, und dazu sind noch die Skalierung von  $\delta E$ ,  $\delta S$ ,  $\delta I$  und  $\delta G$  untereinander stark verschieden.

Die gesuchte Schranke  $\text{diam}(\delta c^\Delta)$  aus Gleichung (4.12) ist geschlossen schlecht abzuschätzen. Eine relativ unscharfe, aber berechenbare, Abschätzung wäre in jedem Falle

$$\text{diam}(\delta c^\Delta) = \max_{i,j} \{|\delta c_i - \delta c_j|\} \leq 2 \max_i \{|\delta c_i|\} = 2\|\delta c^\Delta\|_\infty, \quad (4.15)$$

also das Zweifache des von  $\delta A$  und  $\delta b$  erzeugten komponentenweisen Vorwärtsfehlers in der Lösung von (3.102). Die Literatur liefert Praktische Schranken



für diesen komponentenweisen  $\|\cdot\|_\infty$ -Fehler in  $c^\Delta$  (Higham, 2002, S. 130–131). Da wie eingangs überlegt der numerische Fehler der Lösung  $x$  klein ist im Vergleich zum Einfluss der Strukturfehler  $\delta A$ ,  $\delta b$ , lässt sich die dort anzufindende „naive“ Schranke

$$\|\delta x\|_\infty \leq \| |A^{-1}| |b - A(A + \delta A)^{-1}(b + \delta b)| \|_\infty \quad (4.16)$$

zur Abschätzung verwenden, denn  $\|\delta c^\Delta\|_\infty \leq \|\delta x\|_\infty$ . Hier gehen die Störungen  $\delta A$  und  $\delta b$  nicht normweise ein, sondern werden direkt über das Residuum  $r := b - A(A + \delta A)^{-1}(b + \delta b)$  der Lösung des gestörten Systems berücksichtigt. Die oben besprochenen strukturellen Eigenschaften der Störungen werden damit zumindest nicht vernachlässigt.

Wie scharf ist (4.16) aber und ist die Abschätzung damit überhaupt nützlich? Aus folgenden Gründen lassen sich diese Fragen entweder nicht oder nur negativ beantworten:

1. Abschätzungen der Normen von  $\delta E$ ,  $\delta S$ , etc., sind sehr unscharf, es lässt sich auf ihrer Basis schon keine schärfere Fehlerschranke konstruieren.
2. Die Auswertung der Schranke erfordert die Berechnung des Residuums  $r$ , also schon die Konstruktion des tatsächlichen *und* des gestörten Systems. Allgemein ist nur das gestörte System überhaupt bekannt.
3. Es werden Beträge genommen und daher ist möglicherweise  $\|\delta x\|_\infty \ll \| |A^{-1}| |r| \|_\infty$ .
4. Je nach Verteilung der Fehler ist unter Umständen zusätzlich noch das gesuchte  $\text{diam}(\delta c^\Delta) \ll \|\delta x\|_\infty$ .
5. Die genaue Auswertung erfordert die Berechnung von  $A^{-1}$  oder eine Approximation von  $|A^{-1}| |r|$ , ist also entweder rechnerisch untragbar oder unterschätzt möglicherweise den gesuchten Fehler (Higham, 2002, Kap. 15).

## 4.3 Fehlerverhalten der Oberflächenrekonstruktion

Aus den oben genannten Gründen ist eine geschlossene Abschätzung des möglichen Ausgangsfehlers nicht praktikabel. Eine andere Möglichkeit, eine Idee vom Fehlerverhalten des Verfahrens zu bekommen, ist eine empirische Betrachtung vieler zufällig generierter und dann in kontrollierter Art und Weise

gestörter Einzelsysteme. Die zentrale Frage dabei ist, ob durch Störungen der Eingangsdaten die rekonstruierte Oberfläche dermaßen verzerrt wird, dass die Korrespondenzbildung nicht mehr funktioniert. Stellt sich heraus, dass dennoch ein signifikanter Teil der Oberfläche nutzbar bleibt, ist das trotz fehlender globaler Abschätzung zufriedenstellend. Für die folgenden Untersuchungen wird in diesem Zusammenhang gefordert, dass 50%<sup>2</sup> der Tiefeninformationen der rekonstruierten Oberfläche eine Korrespondenzbildung mit dem Katalog zulassen.

Wie wird vorgegangen? Es macht keinen Sinn, gegebene Daten  $A$  und  $b$  direkt zu stören, denn erstens könnte so  $A + \delta A$  singulär werden, und zweitens korrespondiert ein so zufällig erzeugtes Paar Daten  $A + \delta A$  und  $b + \delta b$  höchstwahrscheinlich zu keinem tatsächlichen aus Kratererkennung resultierenden Satz Daten.

Genausowenig ist eine direkte Störung des Eingangsbildes sinnvoll. Man könnte zwar die Parameter des Kameramodells stören und ein irgendwie geartetes Rauschen auf das Testbild addieren, die Sättigung justieren, etc.. Man hätte dann aber immer noch nur *ein* Testbild und damit auch *ein* System gestört, da die Verteilung der Krater im Bild konstant bliebe. Diese Verteilung bestimmt aber die Struktur des zu lösenden Systems und sollte damit selbst Testparameter sein. Eine Methode zur zufälligen Erzeugung von Testbildern würde die Generierung, Beleuchtung und Abbildung von künstlichen Oberflächen umfassen. Schließlich ließe sich aber selbst so nicht garantieren, dass die hier letztendlich interessante Störung, die der Ellipsenparameter, eine hinreichende Abdeckung hätten.

Daher wird als Ausgangspunkt die direkte Generierung der Ellipsensätze **E** über simple geometrische Modelle von sichtbaren Oberflächen und Addition von kleinen relativen oder absoluten Fehlern auf die einzelnen Ellipsenparameter gewählt. Solche Fehler modellieren die Ausgangsfehler der bildverarbeitenden Verfahrensstufe, der Kratererkennung. Sie ist damit nicht mehr im Testkreis enthalten.

Die Konstruktion der Testdaten verläuft wie folgt: Sei  $S$  eine über  $(u, v)^T$  parametrisierte Oberfläche im Kamera-Koordinatensystem und  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$  ein auf  $S$  mindestens einmal stetig differenzierbares Skalarfeld, dessen Null-Isofläche  $S$  ist (vgl. Abschn. 3.6.1, Def. 3.2). Für geeignet verteilte  $(u_i, v_i)^T$ ,  $i = 1, \dots, N_p$ , ist jeweils

$$p_i = S(u_i, v_i) \in \mathbb{R}^3 \quad (4.17)$$

---

<sup>2</sup>Dies ist keine Beschränkung der Allgemeinheit, die Analyseverfahren funktionieren für jedes solche gewählte Quantil.

ein Punkt auf der Oberfläche in Kamerakoordinaten gegeben. In jedem dieser Punkte beschreibe

$$n_i = \frac{J[F](p_i)}{\|J[F](p_i)\|_2} \quad (4.18)$$

die äußere Einheitsnormale an  $S$ . Mit Kraterradien  $R = (r_1, \dots, r_{N_p})^T$  mit  $r_{\min} \leq r_i \leq r_{\max}$  für sinnvolle  $r_{\min}$  und  $r_{\max}$  lässt sich an jedem Punkt  $p_i$  ein „Kraterrand“

$$C_i = \{q \in \mathbb{R}^3 \mid n_i^T(p_i - q) = 0 \text{ und } \|p_i - q\|_2 = r_i\} \quad (4.19)$$

definieren und die Menge der durch  $S$ ,  $(u_i, v_i)^T$  und  $r_i$  erzeugten Ellipsen ist dann gegeben durch die Projektionen dieser Kraterränder auf die Bildebene:

$$\mathbf{E} = \{E_i = T_C^T(C_i) \mid i = 1, \dots, N_p\} \quad (4.20)$$

Dabei beschreibt  $T_C^T = T_F^T \circ T_C^F$  die perspektivische Projektion auf die Bildebene (vgl. Abschn. 3.6.1, Abb. 3.12). Die Punkte  $p_i$  und die Radien  $r_i$  stellen gleichzeitig den Kraterkatalog für das Testszenario dar:

$$\mathbf{K} := \{(p_i, r_i) \mid i = 1, \dots, N_p\} \quad (4.21)$$

Im Flussdiagramm in Abb. 4.3 ist der obige Vorgang in den Abschnitten „Szenario-Generierung“ und „Messdaten-Generierung“ enthalten.

Die erzeugten Ellipsen  $\mathbf{E}$  stellen für das Testszenario die ungestörten Eingangsdaten für die Splinesystem-Berechnung dar. Jede Ellipse kann in ihren fünf Parametern  $(x, y, a, b, \theta)$  getrennt gestört werden. Diese Störungen lassen sich wie folgt weiter unterteilen bzw. zusammenfassen:

- Die Verschiebung  $(\delta x, \delta y)^T$  des Ellipsenmittelpunktes  $(x, y)^T$  im Bild, gemessen in Pixeln (es wird weiterhin eine Bildbreite von 1024px angenommen).
- Die Stauchung/Streckung  $\delta a$  und  $\delta b$  der Ellipsen-Halbachsen  $a$  und  $b$ , gemessen als Anteil der ungestörten Werte.
- Die Rotation  $\delta \theta$ , um die die Orientierung  $\theta$  der Ellipse gestört wird, gemessen in Radiant.

Möchte man einen Abdeckungstest bezüglich dieser Störparameter implementieren, ergibt sich das ganz praktische Problem der Diskretisierung und Abzählung der zulässigen Störungsmagnituden und -Richtungen. Die nächstliegende Möglichkeit hierzu wäre eine äquidistante Unterteilung dieser Wertebereiche

### 4.3.0 Fehlerverhalten der Oberflächenrekonstruktion

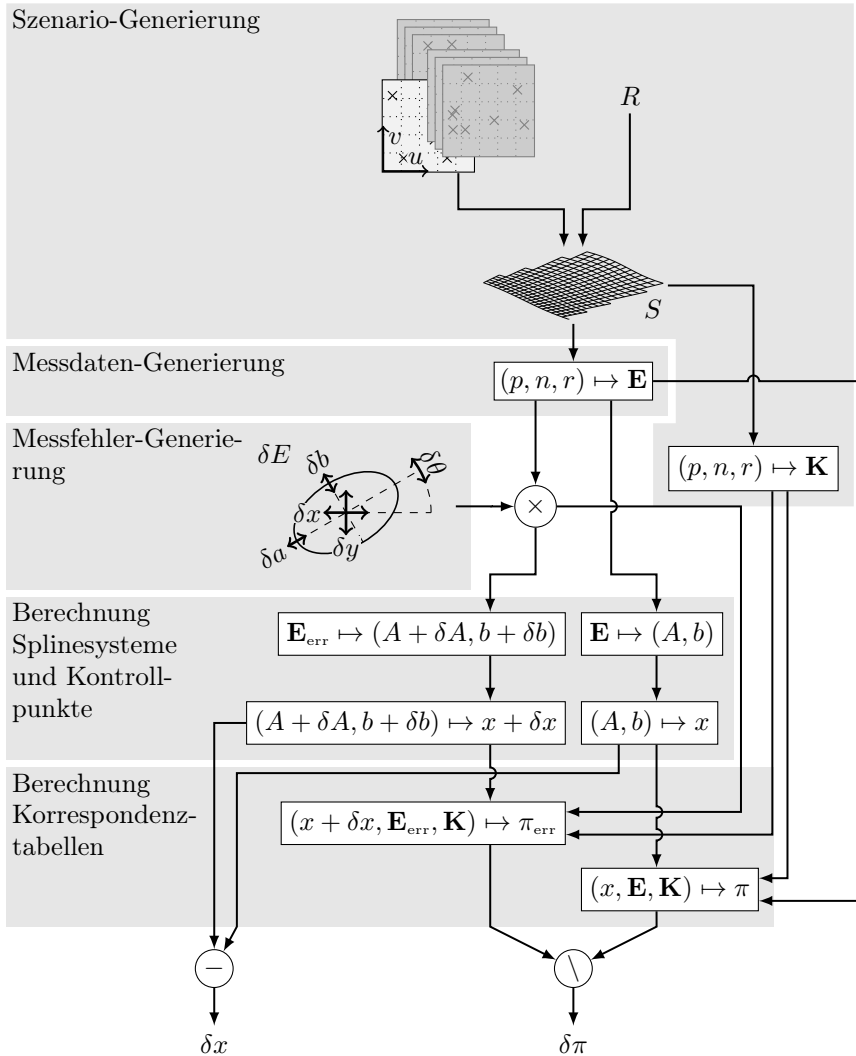


Abbildung 4.3: Systematische Generierung und Störung von Testszenarien mit Datenfluss zwischen den Verfahrensstufen

und Berechnung aller aus den Kombinationen der Störungen resultierenden Systeme. Das wird aber schon für kleine Anzahlen an Testellipsen rechnerisch zu teuer, wie man am Beispiel der Störung des Ellipsenmittelpunktes illustrieren kann: Bei  $N_p = 1$  Ellipse soll der Mittelpunkt  $(x, y)^T$  in zehn Richtung im Bild in drei Schritten bis zu einem festen Betrag  $d_{xy} > 0$  gestört werden. Die Skizze in Abb. 4.4 zeigt, dass hier bereits 30 gestörte Positionen resultieren. Erhöht man die Anzahl der Testellipsen auf  $N_p = 2$ , müsste man nach dem oben vorgeschlagenen Vorgehen die Systeme für alle Kombinationen der Störungen von Ellipse 1 und 2 berechnen. Selbst wenn man Kombinationen von Störungen mit unterschiedlichen Magnituden ausschließt<sup>3</sup>, kommt man pro Störungsbetrag  $d_{xy}$  auf  $10^2$  Kombinationen, also insgesamt  $3 \cdot 10^2 = 300$ . Für  $N_p = 3$  auf  $3 \cdot 10^3$ , usw.

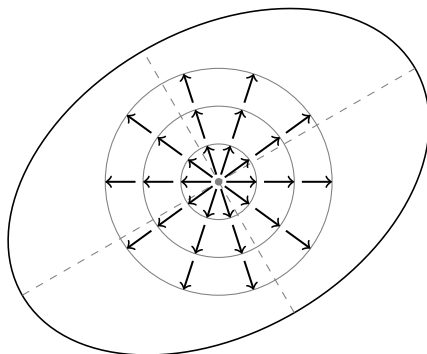


Abbildung 4.4: Diskretisierung der Störung des Ellipsenzentrums

Für die weiteren Parameter  $a$ ,  $b$  und  $\theta$  verhält es sich genauso. In einem realistischen Beispielszenario kann man mit 8 Testkratern pro Szenario rechnen, dies ist eine für den Algorithmus typische Anzahl. Die zulässige Fehlermagnitude muss man in mindestens 3 Schritte unterteilen, um auf mindestens lineares Verhalten schließen zu können. Die Fehlerrichtung der Zentrumsverschiebung sei recht grob in 8 Schritte unterteilt, die der Halbachsen-Stauchung bzw. -Streckung ebenfalls. Die Störung der Rotation nimmt pro Fehlermagnitude ihrer Eindimensionalität wegen nur zwei Werte an. Es ergibt sich eine Bilanz

<sup>3</sup>Dies ist sinnvoll, um die Trennung der Einflüsse von Fehlermagnitude und -Richtung zu gewährleisten.

von

$$\underbrace{3 \cdot 8^8}_{\text{Zentrum}} + \underbrace{3 \cdot 8^8}_{\text{Halbachsen}} + \underbrace{3 \cdot 2^8}_{\text{Rotation}} = 100,664,064 \quad (4.22)$$

zu berechnenden Spline-Systemen. Diese Zahl muss dann noch mit der Anzahl der simulierten Kraterverteilungen und der Anzahl der Unterteilungen des simulierten Betrachtungswinkels multipliziert werden.

Um diesen Aufwand zu reduzieren und auch kleinere Störungen als durch grobe äquidistante Diskretisierung zuzulassen, wird ein Abdeckungstest mit zufallsverteilten Störungen vorgeschlagen. Die verfolgte Strategie ist dann, die *Parameter der Verteilung* der Eingangsstörung zu variieren und in Abhängigkeit von ihnen die Ausgangsfehler zu betrachten.

Die nach dem in Abb. 4.3 dargestellten Testschema zu variierenden Test-szenario-Parameter für die Generierung der ungestörten Testdaten sind:

**Referenzfläche** — Einziger Freier Parameter von  $S$  sei die zentrale Oberflächennormale  $n_S$  in Kamerakoordinaten. Damit liegen alle Testkrater auf einer Ebene, die zentral auf der optischen Achse des Systems in geeignetem Abstand (Ausfüllung des Bildes durch die Testkrater) festgemacht sei. Die Variation von  $n_S$  ist nur in einer Ebene notwendig, da Rotation um die optische Achse äquivalent zur Rotation des Bildes ist:

$$n_S = - \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_n & \sin \varphi_n \\ 0 & -\sin \varphi_n & \cos \varphi_n \end{pmatrix} e_{z,C} \text{ mit } \varphi_n = 0, d_n, \dots, \pi/2 \quad (4.23)$$

$n_S$  wird also äquidistant in Schritten  $d_n$  variiert von Parallelität mit der optischen Achse bis zum rechten Winkel mit ihr. Der Winkel  $\varphi_n$  entspricht damit der Szenariovariable „Kamera-Elevation“.

**Radien der Testkrater** — Die Radien  $R$  der  $N_p$  Testkrater werden aus einer Verteilung gezogen, die eine große Anzahl kleiner und mit wachsendem Radius exponentiell weniger Krater liefert:

$$\frac{1}{N_p} \Sigma(r_i > r) \sim (2r)^{-2.9} \quad (4.24)$$

Diese Verteilung basiert auf Beobachtungen von Shoemaker u. a. anhand von während der Apollo 12-Mission aufgenommenen Daten (Shoemaker u. a., 1970, S. 134). Dies ist eine praktikabel einfache für die Zwecke dieser Tests ausreichende Annäherung an tatsächlich beobachtete Größenverteilungsfrequenzen

auf dem Mond, wenn sie auch gerade für sehr große Krater Abweichungen von der Realität zeigt (Neukum, König und Arkani-Hamed, 1975, S. 226). Die Radien beeinflussen die Splinefläche nicht, wohl aber die Korrespondenzbildung zwischen den erkannten Kratern und dem Katalog.

**Verteilung der Testkrater auf  $S$**  — Die Punkte  $(u, v)^T$  auf  $S$  werden zufällig aus einer Gleichverteilung auf einem Quadrat der Seitenlänge  $w$  um 0 gezogen, für jeweils konstante Radien:

$$\begin{pmatrix} u \\ v \end{pmatrix}_r \sim \mathcal{U} \left( \left[ \pm \frac{w}{2} \right]^2 \right) \quad (4.25)$$

Da  $S$  per Festlegung in Punkt 1) Ebene ist, bleiben die Testkrater auf  $S$  ebenfalls gleichverteilt. Die daraus erzeugten Testellipsen  $E$  sind nach Projektion durch das Kameramodell nicht mehr gleichverteilt im Bild. Die Gesamtmenge der Testpunkte  $(u, v)_R^T$  ist die Vereinigung

$$\begin{pmatrix} u \\ v \end{pmatrix}_R = \bigcup_{r \in R} \begin{pmatrix} u \\ v \end{pmatrix}_r \quad (4.26)$$

der Testpunkte für alle Radien.

Die Fehler, die auf die so generierten Testdaten addiert werden, um gestörte Eingangsdaten für die Splinesystem-Berechnung und die Korrespondenzbildung zu erzeugen, werden wie folgt generiert:

**Störung der Ellipsen-Mittelpunkte** — Die Störungen  $(\delta x, \delta y)^T$  werden aus einer gewickelten Normalverteilung<sup>4</sup> gezogen mit einer Varianz  $\sigma_{xy}^2$  um eine „Hauptstörungsrichtung“  $\mu_{xy}$  als

$$\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = d_{xy} \begin{pmatrix} \cos \varphi_{xy} \\ \sin \varphi_{xy} \end{pmatrix} \text{ mit } \varphi_{xy} \sim \mathcal{N}(\mu_{xy}, \sigma_{xy}). \quad (4.27)$$

Eine Strecke von der Länge  $d_{xy}$  Pixeln verteilt sich also auf die beiden Richtungen.

---

<sup>4</sup>Wahrscheinlichkeitsdichtefunktion  $f_{\mu, \sigma}(\theta) = \frac{1}{\sigma \sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \exp \left( -\frac{(\theta - \mu + 2\pi k)^2}{2\sigma^2} \right)$

**Störung der Ellipsen-Halbachsen** — Die Störungen  $\delta a$  und  $\delta b$  haben den Parameter  $0 < d_{ab} < 1$  der relativen Gesamtänderung und werden mittels eines „Verteilungswinkels“  $\varphi_{ab}$  wieder gewickelt-normalverteilt als

$$\delta a = a d_{ab} \cos(\varphi_{ab}) \text{ und } \delta b = b d_{ab} \sin(\varphi_{ab}) \text{ mit } \varphi_{ab} \sim \mathcal{N}(\mu_{ab}, \sigma_{ab}). \quad (4.28)$$

Hier wird ein relativer Störanteil  $d_{ab}$  (z.B. 5%) auf die beiden Halbachsen verteilt.

**Störung der Ellipsen-Orientierung** — Die Störung  $\delta\theta$  ist eindimensional und wird gleichverteilt aus dem rechtsoffenen Einheitsintervall gezogen, mit dem gewünschten Schiefe-Parameter  $\mu_\theta$  verglichen, und das resultierende Vorzeichen mit festem Betrag  $d_\theta$  multipliziert:

$$\delta\theta = s d_\theta \text{ mit } s = \begin{cases} -1 & : u < \mu_\theta \\ +1 & : u \geq \mu_\theta \end{cases}, \quad u \sim \mathcal{U}([0, 1)). \quad (4.29)$$

Pro generiertem Szenario mit  $N_p$  Testellipsen werden folglich  $N_p$  Fehler aus ihren jeweiligen Verteilungen gezogen, auf die ungestörten Ellipsen **E** addiert und dann in die Verarbeitungskette des Gesamtverfahrens gegeben (Blöcke „Berechnung Splinesysteme und Kontrollpunkte“ und „Berechnung Korrespondenztabelle“ in Abb. 4.3.) Dies wird so oft für jeden ungestörten Ellipsensatz **E** wiederholt, bis die gewünschten Eingangsfehler-Verteilungen hinreichend gut approximiert sind.

„Hinreichend gute“ Approximation ist dann gegeben, wenn ein  $\chi^2$ -Test die vorgegebenen Parameter  $\mu$  (und ggf.  $\sigma$ ) im 95%-Quantil findet. Das beschriebene Vorgehen wird in den Pseudocode-Algorithmen 12 und 13 illustriert.

Diese Teststrategie ist so angelegt, dass die Ausgangsfehler  $\delta x$  und  $\pi_{\text{err}}$  bezüglich der verschiedenen Teil-Eingangsfehler stratifiziert bzw. kategorisiert werden können. Zusammenhänge zwischen Eingangs- und Ausgangsfehlern können erkannt werden und ihr einzelner Einfluss und Beitrag zum Gesamtfehler ist quantifizierbar.

Verschiedene Fehlermodi im Gesamtnavigationssystem (Alignmentfehler der Sensoren, Misskalibration, Vibration, etc.) erzeugen bevorzugt verschiedene der drei Störungskategorien „Ellipsenmittelpunkt“, „Ellipsen-Halbachsen“ und „Ellipsenrotation“, und jede der Kategorien hat wiederum unterschiedlich großen Einfluss auf den Ausgangsfehler  $\delta x$  bzw. auf  $\pi_{\text{err}}$ .

Es wird zunächst daher eine qualitative Betrachtung der Einflüsse der Szenario- und Störungskategorien bzw. -Größen vorgenommen, um einen Eindruck von der Größenordnung des Beitrages jeder einzelnen Unterstörungsart



```

1 Function initialize( $N_{n_S}, N_d, d_{xy}, \mu_{xy}, \sigma_{xy}, d_{ab}, \mu_{ab}, \sigma_{ab}, d_\theta$ )
2    $\mathbb{D}_{n_S} \leftarrow \{n_S = -R_x(\theta)e_{z,\text{cam}} \mid \theta = i\pi/(2N_{n_S}), i = 0, \dots, N_{n_S}\}$ 
3    $\mathbb{D}_R \leftarrow \{r \in \mathbb{R}^+ \text{ nach (4.24)}\}$ 
4    $\mathbb{D}_{d_{xy}} \leftarrow \{d_{xy} = i(d_{xy}^{\max}/N_d) \mid i = 0, \dots, N_d\}$ 
5    $\mathcal{D}_{xy} \leftarrow \mathcal{N}(\mu_{xy}, \sigma_{xy})$ 
6    $\mathbb{D}_{d_{ab}} \leftarrow \{d_{ab} = i(d_{ab}^{\max}/N_d) \mid i = 0, \dots, N_d\}$ 
7    $\mathcal{D}_{ab} \leftarrow \mathcal{N}(\mu_{ab}, \sigma_{ab})$ 
8    $\mathbb{D}_{d_\theta} \leftarrow \{d_\theta = i(d_\theta^{\max}/N_d) \mid i = 0, \dots, N_d\}$ 
9    $\mathcal{D}_\theta \leftarrow \mathcal{U}(\{-1, 1\})$ 
   return :  $\mathbb{D}_{n_S}, \mathbb{D}_R, \mathbb{D}_{d_{xy}}, \mathcal{D}_{xy}, \mathbb{D}_{d_{ab}}, \mathcal{D}_{ab}, \mathbb{D}_{d_\theta}, \mathcal{D}_\theta$ 

```

**Algorithmus 12** : Parameter des Abdeckungstests

zu bekommen und diese mit dem Einfluss der jeweils anderen und dem Einfluss des Absolutbetrages des Fehlers und der Szenariovariable „Kamera-Elevation“ vergleichen zu können. Zu diesem Zweck wird für jede der Fälle  $xy$ ,  $ab$  und  $\theta$  eine Betrachtung der Abhängigkeiten von den Mittelwerten der Störungsrichtungen-Verteilungen zu konstanten Varianzen und danach der Abhängigkeit von den Varianzen zu konstantem Mittelwerten diskutiert werden. Für alle diese Test wird eine grobe äquidistante Unterteilung der Parameterräume der Kamera-Elevation und der absoluten Fehlermagnitudo angenommen. Dies erlaubt, die tatsächlichen Verteilungen der Ausgangsfehler zu bewerten.

Im zweiten Schritt wird dann eine feinere Variation der Eingangsfehlerverteilungen-Mittelwerte zu konstanter Varianz und dann eine Anhebung der Varianz zu jeweils konstanten Mittelwert vorgenommen. Dies ermöglicht eine explorative Modellbildung für Abhängigkeiten repräsentativer Modi der Ausgangsfehlerverteilungen (hier dem Median) von den variierten Eingangsfehlerverteilungen-Modi.

Der letzte Schritt soll dann die gewonnenen Erkenntnisse zu den Abhängigkeiten von den Szenarioparametern zusammenfassen und nutzen in einer Abdeckung der gesamten kombinierten Bandbreite der Verteilungsparameter der Eingangsstörungen.

### 4.3.1 Charakterisierung des Tiefenfehlerverhaltens

Die Berechnung der Spline-Ausgleichsfläche aus gestörten Eingangsdaten  $\mathbf{E}_{\text{err}}$  erzeugt „Tiefenfehler“ in dem Sinne, dass die projektiven Sichtlinien auf die Fläche gestaut oder gelängt werden, und damit die Entfernung zu

```

1 Function test( $N_d, \mu_{xy}, \sigma_{xy}, d_{xy}, \mu_{ab}, \sigma_{ab}, d_{ab}, \mu_{\theta}, d_{\theta}$ )
2    $\mathbb{D}_{n_S}, \mathbb{D}_R, \mathbb{D}_{d_{xy}}, \mathcal{D}_{xy}, \mathbb{D}_{d_{ab}}, \mathcal{D}_{ab}, \mathbb{D}_{d_{\theta}}, \mathcal{D}_{\theta} \leftarrow \text{initialize}(\mathbb{D}_{n_S}, \mathbb{D}_R, \mathbb{D}_{d_{xy}}, \mathcal{D}_{xy}, \mathbb{D}_{d_{ab}}, \mathcal{D}_{ab}, \mathbb{D}_{d_{\theta}}, \mathcal{D}_{\theta})$  /* Algo. 12 */
3    $N_{n_S}, N_d, d_{xy}, \mu_{xy}, \sigma_{xy}, d_{ab}, \mu_{ab}, \sigma_{ab}, \mu_{\theta}, d_{\theta}$ 
4   for  $i = 1, \dots, N_{uv}$  do
5      $P \leftarrow \emptyset$ 
6     for  $r \in \mathbb{D}_R$  do
7        $(u, v) \leftarrow \mathcal{U}([\pm w/2]^2)$ 
8        $P \leftarrow \text{append}((r, u, v))$ 
9     for  $n_S \in \mathbb{D}_{n_S}$  do
10       $\mathbf{C}, \mathbf{K} \leftarrow \emptyset$ 
11      for  $(r, u, v) \in P$  do
12         $p \leftarrow S(n_S; u, v)$ 
13         $\mathbf{K} \leftarrow \text{append}(\mathbf{K}, (p, r))$ 
14         $C \leftarrow \{q \in \mathbb{R}^3 \mid n_S^T(p - q) = 0 \wedge \|p - q\|_2 = r\}$ 
15         $\mathbf{C} \leftarrow \text{append}(\mathbf{C}, C)$ 
16       $\mathbf{E} \leftarrow \emptyset$ 
17      for  $C \in \mathbf{C}$  do
18         $E \leftarrow T_C^1(C)$ 
19         $\mathbf{E} \leftarrow \text{append}(\mathbf{E}, E)$ 
20       $(A, b) \leftarrow \text{surface}(\mathbf{E})$  /* Abschn. 3.7 und 3.8 */
21       $x \leftarrow A^{-1}b$ 
22       $\pi \leftarrow \text{cat\_correspondences}(x, \mathbf{K}, \mathbf{E})$  /* Abschn. 3.9 */
23      for  $\text{var} \in \{xy, ab, \theta\}$  do
24        for  $d_{\text{var}} \in \mathbb{D}_{d_{\text{var}}}$  do
25           $\Delta \mathbf{E}, \mathbf{E}_{\text{err}} \leftarrow \emptyset$ 
26          while  $\Delta \mathbf{E} \not\sim \mathcal{D}_{\text{var}}$  do
27            for  $E \in \mathbf{E}$  do
28               $\delta E \leftarrow \mathcal{D}_{\text{var}}$ 
29               $\mathbf{E}_{\text{err}} \leftarrow \text{append}(\mathbf{E}_{\text{err}}, E + \delta E)$ 
30               $\Delta \mathbf{E} \leftarrow \text{append}(\Delta \mathbf{E}, \delta E)$ 
31             $(A + \delta A, b + \delta b) \leftarrow \text{surface}(\mathbf{E}_{\text{err}})$ 
32             $x + \delta x \leftarrow (A + \delta A)^{-1}(b + \delta b)$ 
33             $\pi_{\text{err}} \leftarrow \text{cat\_correspondences}(x + \delta x, \mathbf{K}, \mathbf{E}_{\text{err}})$ 
34             $\text{err\_data} \leftarrow \text{append}(\text{err\_data}, (\mathbf{E}_{\text{err}}, \pi_{\text{err}}, \delta x))$ 
35  return err_data

```

**Algorithmus 13** : Abdeckungstest

### 4.3.1 Charakterisierung des Tiefenfehlerverhaltens

beobachteten Kratern, die auf dieser Fläche verortet werden, entsprechen variiert. Dies hat Einfluss auf die später folgende Korrespondenzbildung zwischen den beobachteten und den aus dem Katalog bekannten Kratern.

In diesem Abschnitt soll der Einfluss der Störung der Kraterellipsen *qualitativ* bewertet werden. Das heißt in diesem Zusammenhang, dass die Einflüsse der drei oben besprochenen Unterstörungsarten gesondert betrachtet werden sollen, und die dadurch in der Ausgleichsfläche erzeugten relativen Tiefenfehler in Abhängigkeit der Szenarioparameter bewertet werden können.

Das Layout und Elemente der in diesem Abschnitt gezeigten Grafiken der Art der Abb. 4.5 ist in Tabelle 4.2 erläutert.

Tabelle 4.2: Elemente der Fehlergraphen

Element		Beschreibung
unterer Graphenteil	Ordinate links	Tatsächlich erzeugte relative Fehlerverteilung aus Gleichung (4.11) für Kamera-Elevation gemäß unterer Abszisse und absoluter Fehlermagnitude gemäß Abszisse zwischen den Graphenteilen
	Ordinate rechts	Mediane der in Gleichung (4.13) hergeleiteten Schranke mit der Abschätzung von $\text{diam}(\delta x^{\Delta})$ aus Gleichung (4.16) für unabhängige Variablen wie bei linker Ordinate
	Abszisse unten	Elevation der simulierten Kamera, aufsteigend für jeden der unterschiedlich grau unterlegten Blöcke
	Boxplots	<i>Untere/obere Antenne</i> : 2, 5%- und 97, 5%-Quantile der Ausgangsfehlerverteilung, <i>Untere/obere Begrenzung der Box</i> : Erstes und letztes Quartil der Ausgangsfehlerverteilung, <i>Mitteltrenner</i> : Median der Ausgangsfehlerverteilung
	Geraden	Ausgleichsgeraden durch die Mediane der Boxplots gleicher Kamera-Elevation und aufsteigender Fehlermagnitude (erscheinen gekrümmt aufgrund logarithmischer Achsenskalierung)
beide	Abszisse zw. den Graphenteilen	Absolute Größe der Eingangsstörung. Gilt konstant für die Blöcke aufsteigender Kamera-Elevation in den Graphenteilen gleicher Graufärbung.

Tabelle wird fortgesetzt auf folgender Seite

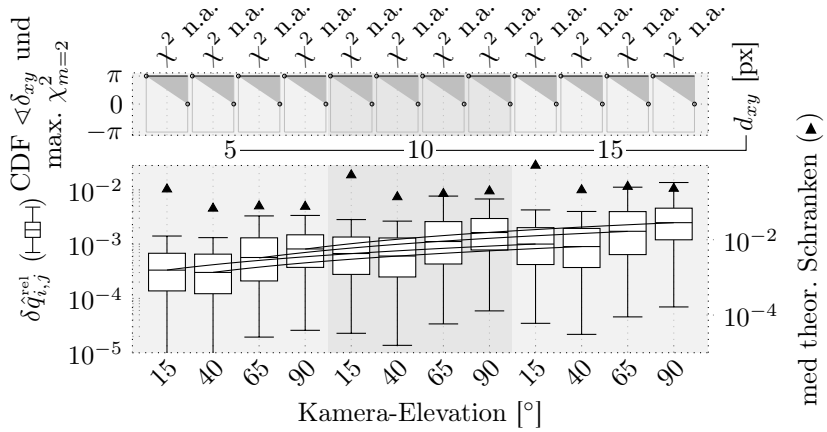
### 4.3.1 Charakterisierung des Tiefenfehlerverhaltens

Fortsetzung Tabelle 4.2

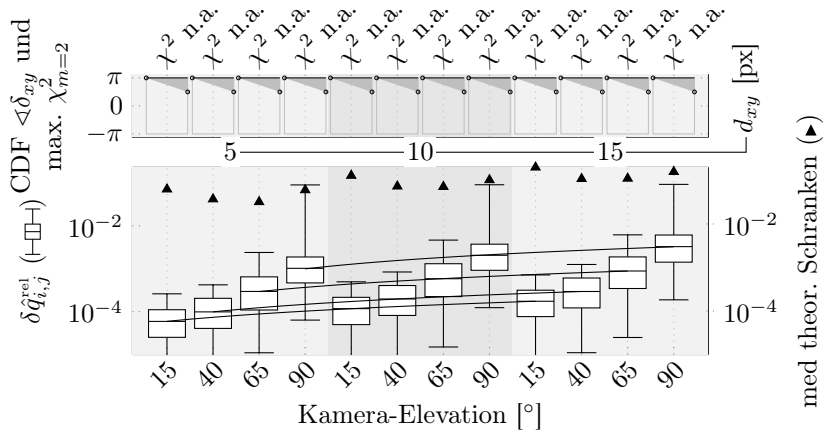
		Element	Beschreibung
oberer Graphenteil		Ordinate	Wertebereich der kumulativen Verteilungsfunktion der Eingangsfehler-Richtungsverteilung (Fälle $xy$ und $ab$ ) bzw. Anteil positiver Vorzeichen soll vs. ist (Fall $\theta$ )
		Abszisse oben	$\chi^2$ -Wert der schlechtesten Approximation der Soll-Eingangsfehler-Richtungsverteilung, die zu dem Verteilungs-Boxplot der betreffenden Spalte beigetragen hat, mit Grad der $\chi^2$ -Verteilung wie jeweils links in der Achsenbeschriftung vermerkt (immer $m = 2$ oder $m = 8$ )
		dunkel gefüllter Bereich	(nur Fälle $xy$ und $ab$ ): Theoretische kumulative Verteilungsfunktion der Soll-Eingangsfehler-Richtungsverteilung
		horizontale Linie	(nur Fall $\theta$ ): Soll-Anteil der Vorzeichen der Ellipsenrotations-Störung mit positivem Umlaufsinn
		Kreise	(nur Fälle $xy$ und $ab$ ): Samples an den verwendeten Intervallgrenzen der tatsächlich realisierten kumulativen Verteilung mit schlechtestem $\chi^2$ -Wert
		gepfeilte Querlinien	(nur Fall $\theta$ ): Tatsächlich realisierte Verteilung der Rotationsvorzeichen mit schlechtestem $\chi^2$ -Wert

**Ellipsenmittelpunkt-Verschiebung: „Fall  $xy$ “** — Zunächst ist der Einfluss der Ellipsen-Mittelpunktverschiebung auf die absoluten Ausgangsfehler  $\delta x$  zu untersuchen. Verursacht werden diese zum Beispiel durch mangelhaftes Alignment der Kamera mit dem Körper-Referenzsystem des Raumfahrzeugs, durch mangelhafte Entzerrung des Bildes, kurz: Durch alle Fehlerquellen, die im Bild lokal hauptsächlich Translation von Pixeln erzeugen.

Die Abb. 4.5 zeigt den Grenzfall von Punktverteilungen der Eingangsfehler-Richtungen von  $\delta x$  und  $\delta y$ . Für diese Verteilung ist der  $\chi^2$ -Test nicht anwendbar (Label „n.a.“ an oberer Abszisse). Ansetzen einer Punktverteilung erzeugt eine Translation aller Ellipsenmittelpunkte  $(x, y)^T$  in die selbe durch  $\mu_{xy}$  bestimmte Richtung. Ein Wert von 0 (wie in Abb. 4.5a) zeigt eine Verschiebung nach unten im Bild an, ein Wert von  $\pi/2$  (wie in Abb. 4.5b) verursacht konstante Verschiebung nach rechts im Bild. An diesen Grafiken lässt sich folgendes beobachten:



(a) Verteilung der relativen Ausgangsfehler für Störungen aller Ellipsenmittelpunkte mit Eingangsfehlern  $\delta x_y \sim \mathcal{N}(0,0)$  und  $d_{xy} \leq 15\text{px}$



(b) Verteilung der relativen Ausgangsfehler für Störungen aller Ellipsenmittelpunkte mit Eingangsfehlern  $\delta x_y \sim \mathcal{N}(\pi/2,0)$  und  $d_{xy} \leq 15\text{px}$

Abbildung 4.5: Variation des Mittelwertparameters der Ellipsen-Mittelpunktstörungsverteilung

### 4.3.1 Charakterisierung des Tiefenfehlerverhaltens

---

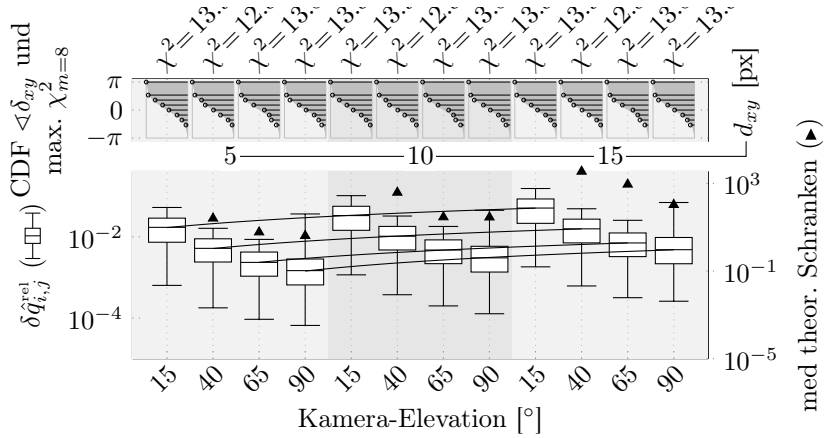
1. Für die Punktverteilungen fallen die mittleren Quartile der Ausgangsfehlerverteilungen vollständig unterhalb der 1-Prozent-Grenze.
2. Bei konstanter Kamera-Elevation lässt sich für alle Verteilungen ein strikt lineares Wachstum der Ausgangsfehler für lineares Wachstum der Eingangsfehlermagnitudo beobachten (Median-Ausgleichsgeraden).
3. Bei konstanter Eingangsfehlermagnitudo und linear wachsender Kamera-Elevation wächst der Ausgangsfehler schneller als linear.
4. Die Mediane der theoretischen Fehlerschranken (schwarze Dreiecke) sind immer größer als die 97,5%-Quantile der empirischen Ausgangsfehlerverteilungen (obere Antennen der Boxplots), meistens sogar um eine ganze Größenordnung. Das belegt leider, dass diese Fehlerschranke, wie schon vermutet, nicht befriedigend scharf ist.

Zur Untersuchung des Einflusses der Streuung der Eingangsfehler-Richtungen wird der Parameter  $\sigma_{xy}$  für festes  $\mu_{xy} = 0$  variiert. Zu der bereits in der Abbildung 4.5a gezeigten Punktverteilung zu  $\sigma_{xy} = 0$  zeigen die Abbildungen 4.6a und 4.6b die Fälle  $\sigma_{xy} = \pi/2$  und  $\sigma_{xy} = \infty$ . Letztere beschreibt dabei praktisch den Fall der unabhängigen Gleichverteilung der Störungsrichtungen im Bild und zeigt damit den zweiten Grenzfall bezüglich der Streuung.

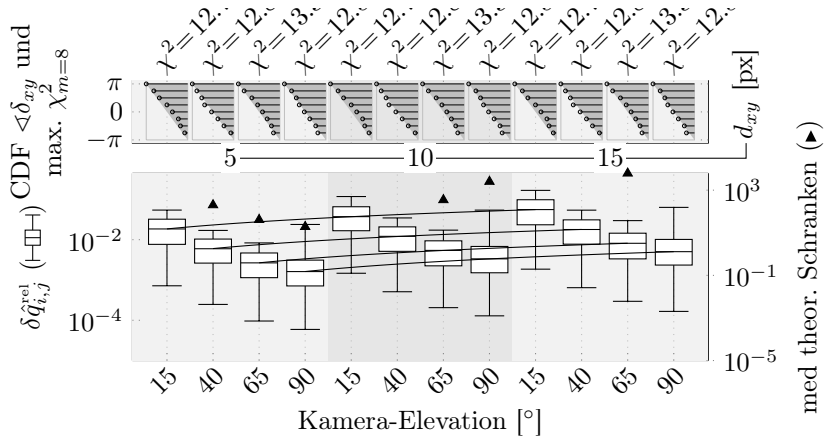
Folgende Beobachtungen lassen sich an den Abbildungen 4.6 machen:

1. In Abb. 4.5 wachsen die Ausgangsfehlerverteilungen mit wachsender Kamera-Elevation. Hier, bei nichtverschwindenden Streuungen der Eingangsfehler-Richtungen, werden sie dagegen kleiner, je weiter sich der Blickwinkel dem Nadir annähert. Dieses Verhalten überrascht nicht, wenn man bedenkt, dass genau im Nadirblickenden Fall die Varianz aller möglichen Distanzen zur Oberfläche am geringsten ist.
2. Die Ausgangsfehlerverteilungen wachsen insgesamt kaum noch beim Übergang zur Gleichverteilung. Die Eingangsstörungen mitteln sich also heraus.
3. Das streng lineare Wachstum der Ausgangsfehler über lineares Eingangsfehlerwachstum besteht weiter.
4. Die theoretischen Fehlerschranken sind für alle nichtverschwindenden Streuungen völlig nutzlos (teilweise jenseits  $10^4$  und nicht mehr dargestellt).

Die Konstruktion der Referenz-Ausgleichsfläche ist bezüglich zufälliger Störung der Ellipsenmittelpunkt-Positionen mit bereits unrealistisch groß angesetzten



(a) Verteilung der relativen Ausgangsfehler für Störungen aller Ellipsenmittelpunkte mit Eingangsfehlern  $\delta_{xy} \sim \mathcal{N}(0, \pi/2)$  und  $d_{xy} \leq 15\text{px}$



(b) Verteilung der relativen Ausgangsfehler für Störungen aller Ellipsenmittelpunkte mit Eingangsfehlern  $\delta_{xy} \sim \mathcal{N}(0, \infty)$  und  $d_{xy} \leq 15\text{px}$

Abbildung 4.6: Variation des Varianzparameters der Ellipsen-Mittelpunktstörungsverteilung

### 4.3.1 Charakterisierung des Tiefenfehlerverhaltens

---

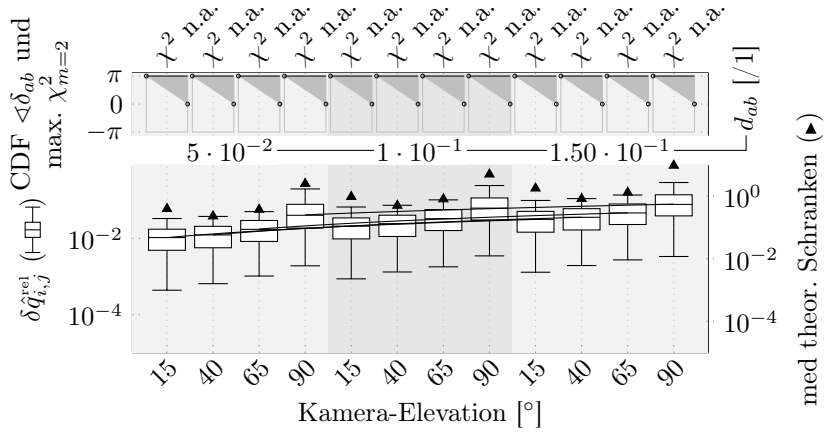
Störbeträgen nicht übermäßig empfindlich. Die oberen Quantile streuen bis auf den nadirblickenden Fall nicht erheblich oberhalb des Medians der Ausgangsfehlerverteilungen.

**Ellipsen-Halbachsenstreckung: „Fall  $ab$ “** — Analog wird jetzt die Wirkung relativer Störung der Ellipsenhalbachsen auf die Berechnung der Ausgleichsfläche untersucht. Die Störstrategie besteht darin, die Halbachsen  $a$  und  $b$  um einen linear steigenden Gesamtprozentanteil  $d_{ab}$  zu strecken, der wieder durch eine Störungs-„Richtung“  $\angle\delta_{ab}$  auf die beiden Halbachsen verteilt wird gemäß  $\delta_a = d_{ab} \cos(\angle\delta_{ab})$  und  $\delta_b = d_{ab} \sin(\angle\delta_{ab})$ .

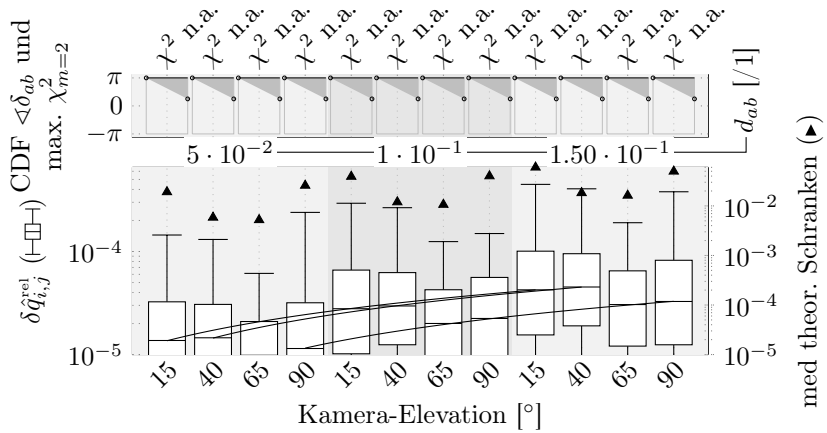
Die Graphen 4.7 zeigen zunächst wieder den Fall rotierender Punktverteilungen. Dabei erzeugt der Fall  $\mu_{ab} = 0$  (Abb. 4.7a) Störung ausschließlich in  $a$  und der Fall  $\mu_{ab} = \pi/4$  (Abb. 4.7b) gleichmäßig in  $a$  und  $b$ . An diesen Graphen lässt sich folgendes feststellen:

1. Werden alle beobachteten Kraterellipsen mit selber Magnitude in nur der großen Halbachse  $a$  gestreckt, ist der Effekt auf die Berechnung der Ausgleichsfläche ähnlich groß wie bei der Mittelpunktverschiebung der Ellipsen.
2. Bei  $\mu_{ab} = \pi/4$  streuen die Ausgangsfehler robust unterhalb der festgelegten Signifikanzgrenze von  $10^{-4}$  und verschwinden großenteils sogar unter der Darstellungsgrenze  $10^{-5}$ . Das ist deswegen nicht überraschend, da in diesem Falle alle Ellipsen konstant skaliert werden, ohne sie zu verzerren. Die Ellipsengröße hat aber keinen Einfluss auf das Ergebnis, da sie die rekonstruierte Oberflächennormale nicht ändert. Die restlichen Ausgangsfehler sind dem Abtastverhalten der Eingangsfehlerverteilungen geschuldet.
3. Die theoretischen Fehlerschranken sind für den Fall  $ab$  etwas schärfer als im vorangegangenen Fall  $xy$ , jedoch nicht erheblich oder in einem Maße, das sie nützlich machen würde.
4. Die Ausgangsfehler streuen in der ungünstigeren Verteilung noch um den 10%-Bereich.
5. Für linear steigende Eingangsfehlermagnitude besteht überall ein robust linearer Zusammenhang mit der Verteilungsparameter der Ausgangsfehler.
6. Mit steigender Kamera-Elevation wachsen die Ausgangsfehler. Das lässt sich daher erklären, dass für kleine Elevationen die Varianz der Parameter





(a) Verteilung der relativen Ausgangsfehler für Störungen aller Ellipsenhalbachsen mit Eingangsfehlern  $\delta_{ab} \sim \mathcal{N}(0, 0)$  und  $d_{ab} \leq 15\%$



(b) Verteilung der relativen Ausgangsfehler für Störungen aller Ellipsenhalbachsen mit Eingangsfehlern  $\delta_{ab} \sim \mathcal{N}(\pi/4, 0)$  und  $d_{ab} \leq 15\%$

Abbildung 4.7: Variation des Mittelwertparameters der Ellipsen-Halbachsen-störungsverteilung

### 4.3.1 Charakterisierung des Tiefenfehlerverhaltens

---

der beobachteten Ellipsen kleiner ist als im nadirblickenden Szenario, wo der Nadir selbst im Bild ist.

Da sich aus der Auflösung eines Kamerabildes eine natürliche untere Schranke für die Fehler der Ellipsenerkennung herleiten lässt, sind kleinere im Bild erkannte Krater im Verhältnis deutlich größeren Störungen der Halbachsen im hier vorgestellten Sinne unterworfen. Zum Vergleich: Um einen im Bild 100 Pixel durchmessenden Krater um 15% zu strecken, muss ein Fitting-Fehler von 15 Pixeln auftreten, was im höchsten Maße unwahrscheinlich ist. Dagegen muss der Ellipsenfit bei einem 10 Pixel durchmessender Krater nur um 1.5 Pixel fehlerhaft sein, um die selbe relative Störung zu erzeugen, was von der Größenordnung der Bildauflösung ist und daher regelmäßig vorkommt. Dies begründet die Präferenz für große Krater in der Auswahlstufe der Ankerpunkte für die Ausgleichsflächen-Berechnung.

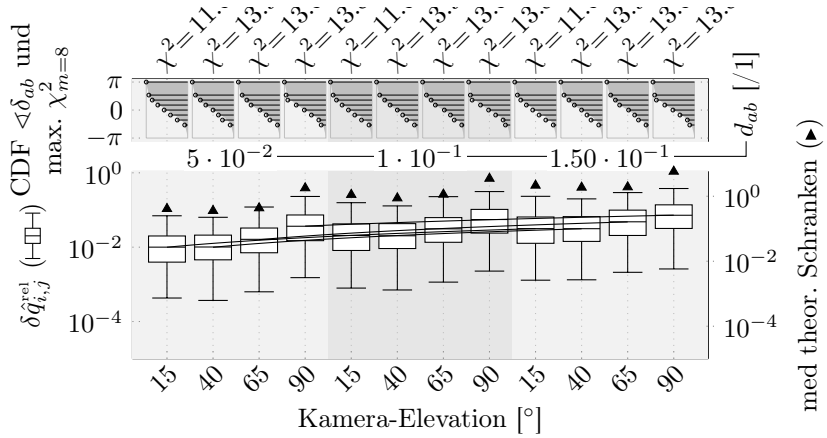
Fehler der in diesen Graphen gezeigten Art sind neben den algorithmischen Fehlern in realistischen Anwendungen beispielsweise Folge von Bildscherungen, etwa durch Alignmentfehler der Bildgebenden Sensoren, die einen nicht-orthogonalen Durchstoß der optischen Achse durch den Sensor verursachen. Solche mechanischen Mängel werden aber im Allgemeinen niemals die hier getesteten Aufmaße von 15% Scherung erreichen.

Nach Variation von  $\mu_{ab}$  sind in Abb. 4.8 jetzt die Ergebnisse der Variation des Streuungsparameters  $\sigma_{ab}$  aufgetragen mit den Werten  $\sigma_{ab} = \pi/2$  bzw.  $\sigma_{ab} = \infty$ , zusätzlich zum in Abb. 4.7a bereits gezeigten Fall  $\sigma_{ab} = 0$ .

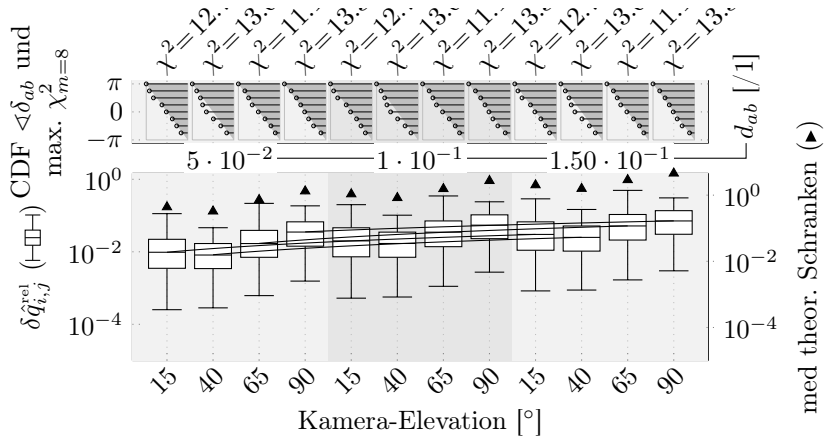
Im Vergleich mit den Abbildungen zum Vergleich der Punktverteilungen und darüber hinaus zeigt sich, dass, anders als im Fall  $xy$ , sich die Ausgangsfehlerverteilungen mit wachsendem  $\sigma_{ab}$  nicht signifikant verschieben, sondern nur etwas breiter streuen. Im Median lässt sich weiterhin derselbe Ausgangsfehler erwarten.

**Ellipsenrotationsstörung: „Fall  $\theta$ “** — Die letzte Störungsvariable ist die Rotation der Ellipse, gemessen als Winkel  $\theta$  zwischen positiver erster Bildkoordinate (Richtung ansteigende Zeilen nach unten) und der großen Halbachse  $a$  der jeweiligen Ellipse. Die Magnitude  $d_\theta$  der Störung beträgt bis zu  $15^\circ$ . Anders als in den Fällen  $xy$  und  $ab$  ist die zu störende Variable eindimensional und ihre Eingangsfehlerverteilung hat deswegen nur den Parameter  $\mu_\theta$  aus Gleichung (4.29). Der Anteil der rechts- vs. linksherum mit festem Störwinkel  $d_\theta$  gerichteten Störung ist dadurch festlegt.

Die Eingangsfehlermagnitude ist hier eine absolute Störung des Ellipsenwinkels und es wird wieder der Fall  $0^\circ$  Grad (der Referenzfall) mit dem



(a) Verteilung der relativen Ausgangsfehler für Störungen aller Ellipsenhalbachsen mit Eingangsfehlern  $\delta_{ab} \sim \mathcal{N}(0, \pi/2)$  und  $d_{ab} \leq 15\%$



(b) Verteilung der relativen Ausgangsfehler für Störungen aller Ellipsenhalbachsen mit Eingangsfehlern  $\delta_{ab} \sim \mathcal{N}(0, \infty)$  und  $d_{ab} \leq 15\%$

Abbildung 4.8: Variation des Varianzparameters der Ellipsen-Halbachsenstörungsverteilung

### 4.3.1 Charakterisierung des Tiefenfehlerverhaltens

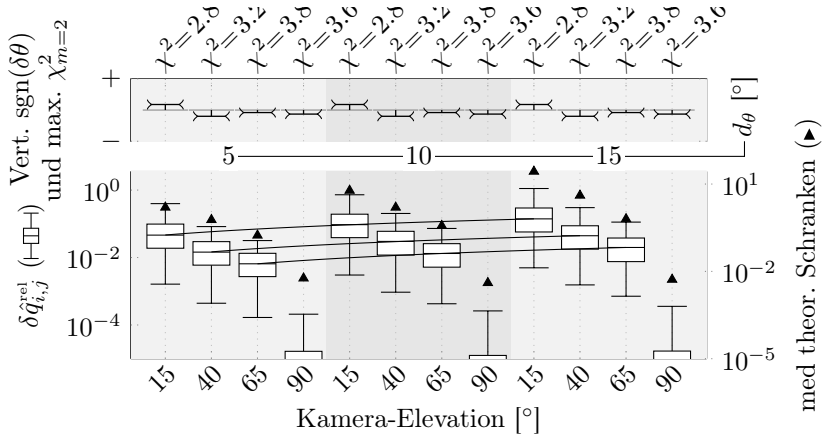
---

gestörten Fall  $15^\circ$  verglichen.

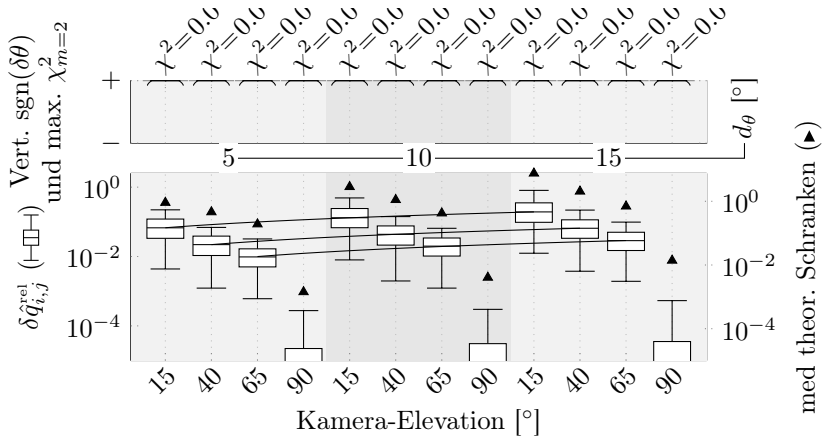
Abbildung 4.9a zeigt den Fall  $\mu_\theta = 1/2$ . Hier werden gleichviele Ellipsen rechts- wie linksherum rotiert. Abbildung 4.9b zeigt die Ausgangsfehler, die resultieren, wenn alle Ellipsen mit konstantem Störwinkel  $d_\theta$  im positiven Drehsinn rotiert werden.

Das gezeigte Ausgangsfehlerverhalten über diese Variation von  $\mu_\theta$  lässt sich wie folgt zusammenfassen:

1. Die Variation von  $\mu_\theta$  hat kaum Einfluss auf die Ausgangsfehlerverteilungen.
2. Die Ausgangsfehler werden mit steigender Kamera-Elevation kleiner. Im nadirblickenden Fall sind die Fehler sogar vernachlässigbar klein. Das ist so auch zu erwarten, da die Winkeländerung in der von einer Ellipse abgeleiteten Oberflächennormale immer kleiner wird, je mehr sich die Ellipse dem Kreis annähert. Die Normalen stellen Koeffizienten des Ausgleichssystems der rekonstruierten Oberfläche dar, kleine Änderung hier bedeutet also kleine Änderung in der Ausgleichslösung.
3. Die Parameter der Ausgangsfehlerverteilung verhalten sich wie schon in den Fällen  $xy$  und  $ab$  linear über die Eingangsfehlermagnituden.
4. Die Mediane der theoretischen Fehlerschranken bewegen sich wieder bis zu einer Größenordnung über den empirischen Ergebnissen. Sie sind also auch in diesem Zusammenhang nicht nützlich.
5. Die Ausgangsfehlerverteilungen streuen etwas höher als in den Tests der anderen Störvariablen beobachtet wurden. Der Algorithmus ist also am empfindlichsten gegenüber dieser Art von Eingangsfehler.



(a) Störende Rotation aller Ellipsen mit mittlerem Anteil 50% im positiven Drehsinn und konstantem Störwinkel  $d_\theta \leq 15^\circ$



(b) Störende Rotation aller Ellipsen 100% im positiven Drehsinn mit konstantem Störwinkel  $d_\theta \leq 15^\circ$

Abbildung 4.9: Variation der mittleren Richtungen der Stör-Rotationen der Ellipsen

#### 4.3.2 Linienschnitte des Eingangsfehlerraumes

Im vorangehenden Unterabschnitt wurden die Ausgangsfehlerverteilungen grob charakterisiert, in Abhängigkeit von Einzelwerten der Eingangsfehler-Richtung und -Streuung. Bei der folgenden vollen Abdeckung des Raumes dieser Richtungen und Streuungen können die vollen Ausgangsfehlerverteilungen nicht mehr dargestellt werden, daher wird repräsentativ jeweils ihr Median verwendet. Da durch die oben besprochenen Mechanismen zu jedem Verteilungsparametersatz sichergestellt ist, dass die gewünschte Verteilung hinreichend gut ( $\chi^2$ ) realisiert wird, kann man den Parameterraum äquidistant abtasten und daraufhin auf die Suche nach funktionalen Zusammenhängen zwischen Eingangsstörung und Ausgangsfehler gehen. Die Variation der Kamera-Elevation und des Eingangsfehler-Absolutbetrages geschehen in den selben Schritten wie in der oben gemachten Charakterisierung.

**Ellipsenmittelpunkt-Verschiebung: „Fall  $xy$ “** — Abbildung 4.10 zeigt im linken oberen Graphen das Verhalten der Mediane der Ausgangsfehlerverteilungen, abgetragen über die Mittelwerte  $-\pi < \mu_{xy} < \pi$  der Eingangsfehlerverteilungen. Die Streuung  $\sigma_{xy}$  wird zunächst 0 gehalten, es handelt sich also um Punktverteilungen.

Stratifiziert man diese Ausgangsfehlermediane nach konstanter Kamera-Elevation und konstantem Eingangsfehlerbetrag  $d_{xy}$ , erhält man die Kurven des linken mittleren Graphen. Dort lassen sich gut periodische Verläufe über  $\mu_{xy}$  erkennen, die Kurven lassen sich jeweils gut approximieren durch die eingezeichneten Modelle mit Parametern  $p_1$  und  $p_2$  ermittelt durch nichtlineare Regression gegen die nach konstanter Kamera-Elevation stratifizierten Daten. Der untere linke Graph zeigt die Residuen der Daten mit ihren jeweils zugehörigen Modellen. Man erkennt Restabhängigkeiten halber Periode, die Residuen sind aber schon im Bereich von  $10^{-4}$ , eine weitere Analyse lohnt sich hier nicht.

Der periodische Zusammenhang bricht bei nadirblickender Kamera zusammen und geht in eine Konstante über. Dies war zu erwarten, da in diesem Fall die Normale der simulierten Oberfläche genau in die Kamera zeigt und daher alle Verschiebungen der Ellipsen-Mittelpunkte ihr gegenüber symmetrisch sind.

Die Graphen der rechten Seite zeigen die Modellparameter  $p_1$  und  $p_2$ , stratifiziert nach der Kamera-Elevation und abgetragen über den Eingangsfehlerbetrag. Es lassen sich bei beiden Parametern starke lineare Abhängigkeiten vom Eingangsfehlerbetrag ablesen. Mit der modellierten trigonometrischen

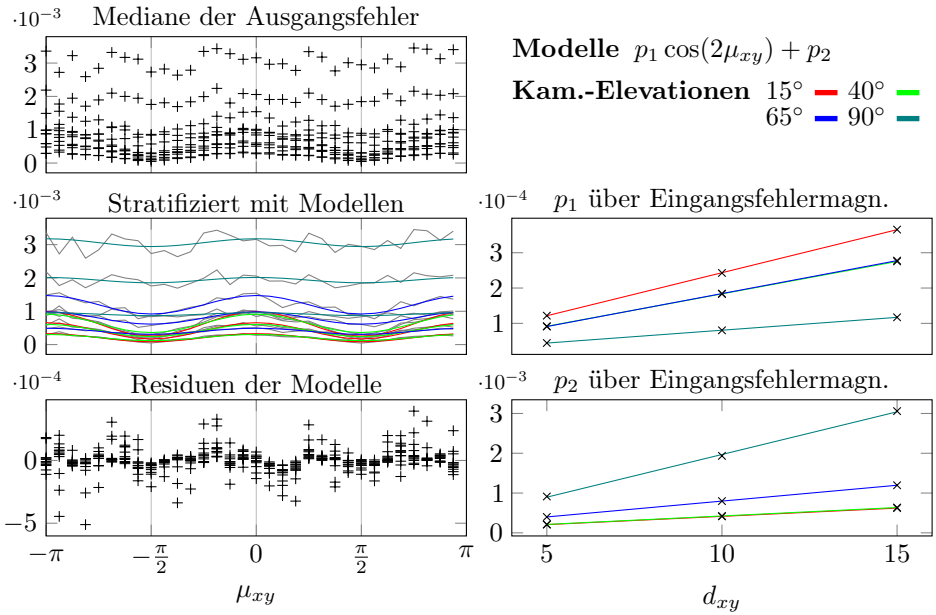


Abbildung 4.10: Nichtlineare Modellierung der Mediane der Ausgangsfehlerverteilungen für rotierende Punktverteilung der Eingangsfehlerstreuung der Ellipsenmittelpunkte.

Abhängigkeit vom Eingangsfehler-Verteilungsmittelwert und diesen Parametern lassen sich vom Ausgangsfehler von der Größenordnung ca.  $3 \cdot 10^{-3}$  bereits alles bis auf ca.  $2 \cdot 10^{-4}$  erklären, also  $\approx 93\%$ .

Abb. 4.11 zeigt das Ausgangsfehlerverhalten für Variation des Eingangsfehlerverteilungs-Parameters  $\sigma_{xy}$  zwischen  $10^{-5}$  und 50 für konstantes  $\mu_{xy} = 0$ . Der erste Wert ist ein minimaler Schritt aufwärts von der in Abb. 4.10 behandelten Punktverteilung und  $\sigma_{xy} = 50$  approximiert wieder die Gleichverteilung der Eingangsfehlerrichtungen. Es finden sich die Maxima der Mediane der Ausgangsfehlerverteilungen konsistent bei der maximalen Streuung der Eingangsstörungsrichtungen.

Stratifiziert man erneut die Ausgangsfehlermediane nach konstanter Kamera-Elevation, lässt sich eine logarithmische Abhängigkeit vom Modell im rechten oberen Kasten in der Abbildung ablesen. Die Residuen der jeweils zugehörigen Modelle sind links unten gezeigt, sie sind wieder vernachlässigbar

### 4.3.2 Linienschnitte des Eingangsfehlerraumes

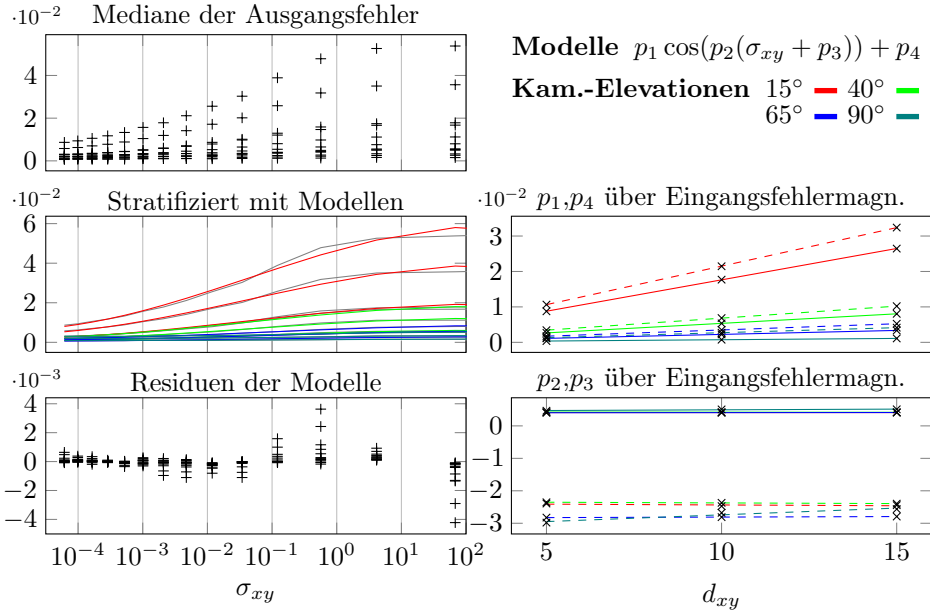


Abbildung 4.11: Nichtlineare Modellierung der Mediane der Ausgangsfehlerverteilungen für wachsende Standardabweichung der Fehler der Ellipsenmittelpunkte (jeweils erster Parameter pro Graph in durchgezogener Linie).

klein. Es lässt sich also wiederum der Großteil der Ausgangsfehler allein durch den modellierten Zusammenhang erklären. Mitte-rechts sind die Parameter  $p_1$  und  $p_4$  der Amplitude bzw. des vertikalen Versatzes des Modells abgetragen mit besten Ausgleichsgeraden. Man kann wieder einen starken linearen Zusammenhang der Parameter mit dem Eingangsfehlerbetrag ablesen. Der Vollständigkeit halber sind rechts unten die Parameter  $p_2$  und  $p_3$  gezeigt, sie sind aber mehr oder weniger konstant.

**Ellipsen-Halbachsenstreckung: „Fall  $ab$ “** — Abbildung 4.12 illustriert das analoge Vorgehen für die Störung der Ellipsen-Halbachsen. Man erkennt im linken oberen Teilgraphen, dass die Ausgangsfehlermediane teilweise bei 10% liegen für  $\mu_{ab}$  bei  $-\pi/4$  bzw.  $3/4\pi$ . Das sind genau die Punkte, in denen die relative Störung einer Ellipsen-Halbachse Null und die der anderen maximal, bzw. umgekehrt, ist. An den Stellen  $\pi/4$  und  $-3/4\pi$  dagegen sind



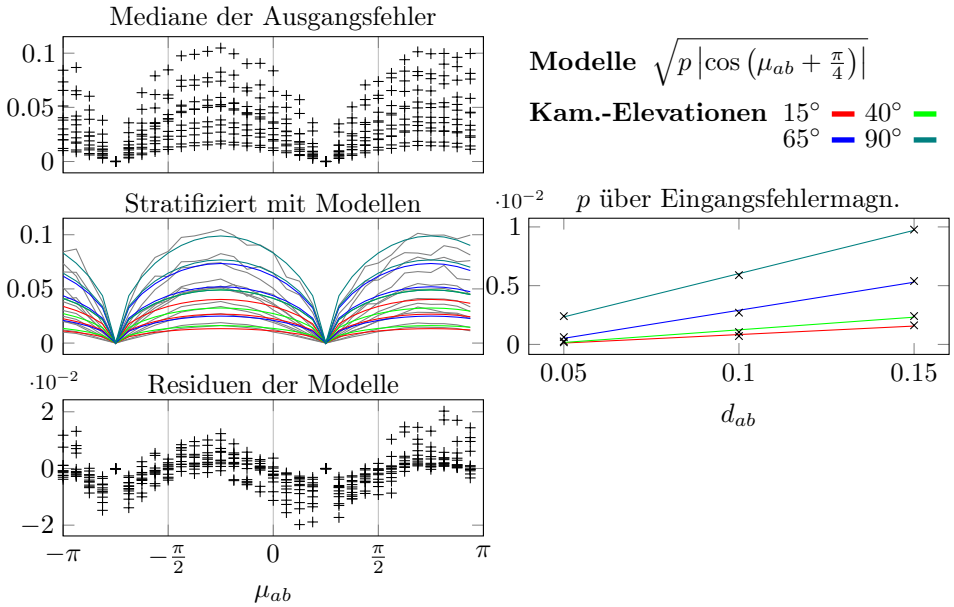


Abbildung 4.12: Nichtlineare Modellierung der Mediane der Ausgangsfehlerverteilungen für rotierende Punktverteilung der Eingangsfehlerstreuung der Ellipsenhalbachsen.

die relativen Störungen identisch und es liegt eine reine Skalierung der Ellipse vor. Die zu diesen speziellen Punkten in Unterabschnitt 4.3.1 zum Fall  $ab$  gemachten Aussagen werden dadurch bestätigt.

Werden die Ausgangsfehlermediane wieder stratifiziert nach konstanter Kamera-Elevation und konstantem Eingangsfehlerbetrag, lassen sich erneut regelmäßige Kurven erkennen. Die Modelle über diesmal nur einen Parameter  $p$  (siehe rechter oberer Kasten in der Abbildung) erklären wiederum genug von der Streuung ( $\approx 85\%$ ), der unmodellerte Restfehler liegt bei ca. 1.5%. Es verbleibt eine sichtbare periodische Abhängigkeit von  $\mu_{ab}$  durch mangelnde Modelltreue. Die lineare Abhängigkeit des Parameters  $p$  vom Eingangsfehlerbetrag lässt sich im Graphen auf der rechten Seite ablesen, wiederum stratifiziert nach der Kamera-Elevation.

Die dazugehörige Abb. 4.13 illustriert wieder die Abhängigkeit der Ausgangsfehler von der Streuung  $\sigma_{ab}$  der Eingangsfehlerrichtungen um  $\mu_{ab} = 0$ .

### 4.3.2 Linienschnitte des Eingangsfehlerraumes

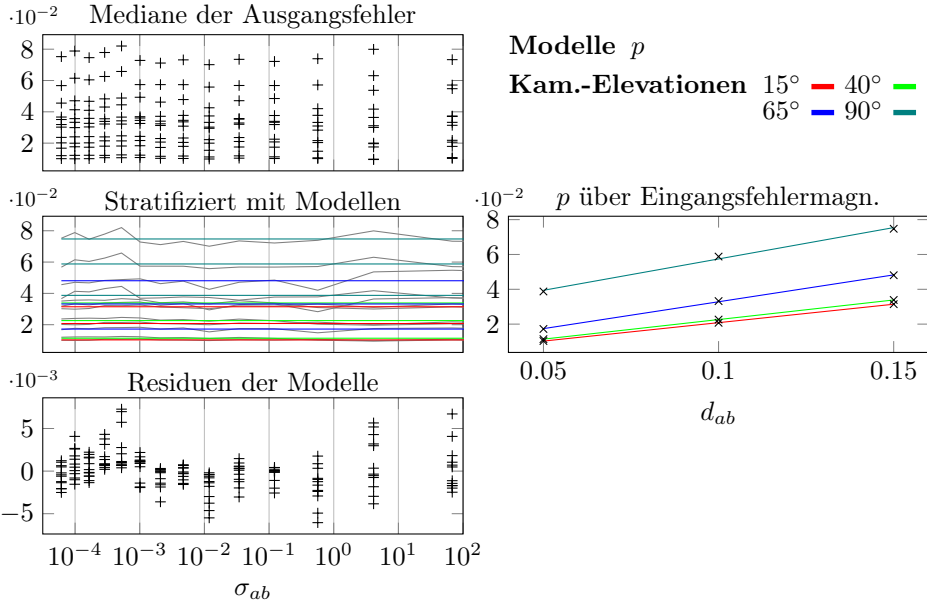


Abbildung 4.13: Konstante Modellierung der Mediane der Ausgangsfehlerverteilungen für wachsende Standardabweichung der Fehler der Ellipsenhalfachsen.

Es ist auffällig, dass die Mediane der Ausgangsfehlerverteilungen offenbar nicht mit  $\sigma_{ab}$  wachsen. Die besten Modelle, die sich hier finden lassen, sind konstant. Sie erklären von den ursprünglich maximalen ca. 8% Fehlern gute 94%, die unmodellierten Restfehler bewegen sich größtenteils unterhalb eines halben Prozentes. Die Konstanten  $p$  zeigen erneut eine Abhängigkeit von der Kamera-Elevation.

**Ellipsenrotationsstörung: „Fall  $\theta$ “** — Die Abbildung 4.14 vervollständigt die Analyse der Abhängigkeit von den Eingangsfehlerrichtungen. Die Ausgangsfehlermediane zeigen hier außer bei kleinsten Kamera-Elevationen kaum Abhängigkeit von  $\mu_\theta$ . Zur Modellierung werden daher Konstanten  $p_\theta$  angesetzt, die allerdings eine leichte periodische Abhängigkeit ignorieren. Wie man im linken unteren Teilgraphen sieht, erklären sie dennoch den Großteil (ca. 95%) der ursprünglich maximal ca. 20% Ausgangsfehler.

Die Stratifizierung der Parameter  $p_\theta$  nach Kamera-Elevation und Betrachtung

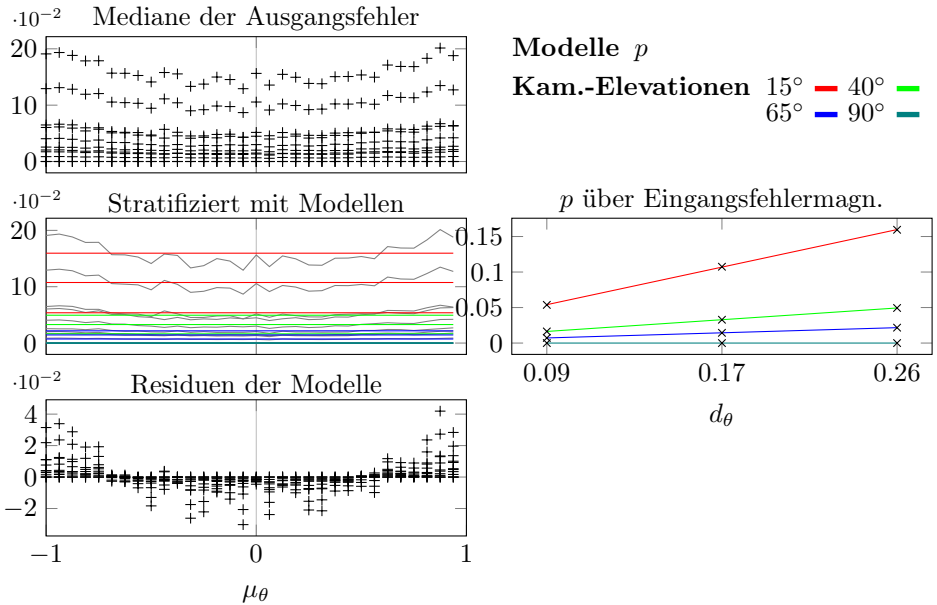


Abbildung 4.14: Konstante Modellierung der Mediane der Ausgangsfehlerverteilungen für wandernde Schiefe der Rotations-Eingangsfehlerrichtungen.

tung über den Eingangsfehlerbetrag zeigt dasselbe lineare Verhalten wie in den vorangegangenen Analysen.

### 4.3.3 Vollabdeckung des Eingangsfehlerraumes

Die oben gemachten Analysen der Abhängigkeiten der Ausgangsfehlergrößen von den Parametern der Eingangsfehler-Richtungsverteilungen zeigen, dass konsistent um 90% der Ausgangsfehler erklärt werden können durch Wissen über den maximalen Betrag der Eingangsfehler (Verschiebungen/Stauch-Streckungen/Rotationen) und der Szenariovariable Kamera-Elevation gegenüber der beobachteten Oberfläche. Die Analyse der Rotationsparameter-Störung ist damit bereits abgeschlossen, denn sie ist eindimensional. Für die Störungen der Ellipsenmittelpunkte und der -Halbachsen muss noch das Fehlerverhalten einer gemischten ( $\mu$  und  $\sigma$ ) Abdeckung überprüft werden.

**Ellipsenmittelpunkt-Verschiebung: „Fall  $xy$ “** — Es muss ausgeschlossen werden, dass die Ausgangsfehler bei den bis jetzt getesteten  $\mu_{xy}$  und  $\sigma_{xy}$  ein Minimum aufweisen und daher von einem nur daraus abgeleiteten Vorhersagemodell unterschätzt werden könnten. Daher werden jetzt gleichzeitig die im Unterabschnitt 4.3.2 benutzen jeweiligen Definitionsbereiche abgedeckt und wieder der relative Fehler der Tiefe aufgetragen. Man erhält durch Stratifizierung nach dem Eingangsfehler-Absolutbetrag und der Kamera-Elevation die Fehlerflächen in Abb. 4.15.

Die für konstantes  $\sigma_{xy}$  in Abb. 4.10 benutzte Modellfunktion war (mit der Variable entsprechend indizierten Parametern)

$$p_{1,\mu_{xy}} \cos(2\mu_{xy}) + p_{2,\mu_{xy}} \quad (4.30)$$

Man erkennt, dass der Amplitudenparameter  $p_1$  des Modells der Ausgangsfehlermediane über  $\mu_{xy}$  bei  $\sigma_{xy} = 0$  am größten ist in Abb. 4.15 und für wachsende  $\sigma_{xy}$  kleiner wird. Am deutlichsten ist dies bei kleinen Kamera-Elevationen zu beobachten (oberste Graphen), da bei nadirblickender Kamera die Ausgangsfehler klein genug werden, um vom Rauschen dominiert zu werden (unterste Graphen).

Der Verlauf der Ausgangsfehlermediane für wachsende  $\sigma_{xy}$  entspricht qualitativ bei allen konstanten  $\mu_{xy}$ -Linien der in Abb. 4.11 gefundenen Modellfunktion

$$p_{1,\sigma_{xy}} \cos(p_{2,\sigma_{xy}}(\sigma_{xy} + p_{3,\sigma_{xy}})) + p_{4,\sigma_{xy}} \quad (4.31)$$

für logarithmische  $\sigma_{xy}$ -Skala. Die Fehlermedian-Flächen in der Abb. 4.15 zeigen, dass die Ausgangsfehler für wachsende  $\sigma_{xy}$  erheblich stärker wachsen, als in Richtung der  $\mu_{xy}$ . Vernachlässigt man die Variation in dieser Richtung und korrigiert man die empirischen Werte nur um das konstant über  $\mu_{xy}$  fortgesetzte Modell (4.31), erhält man bereits eine verlässliche Vorhersage.

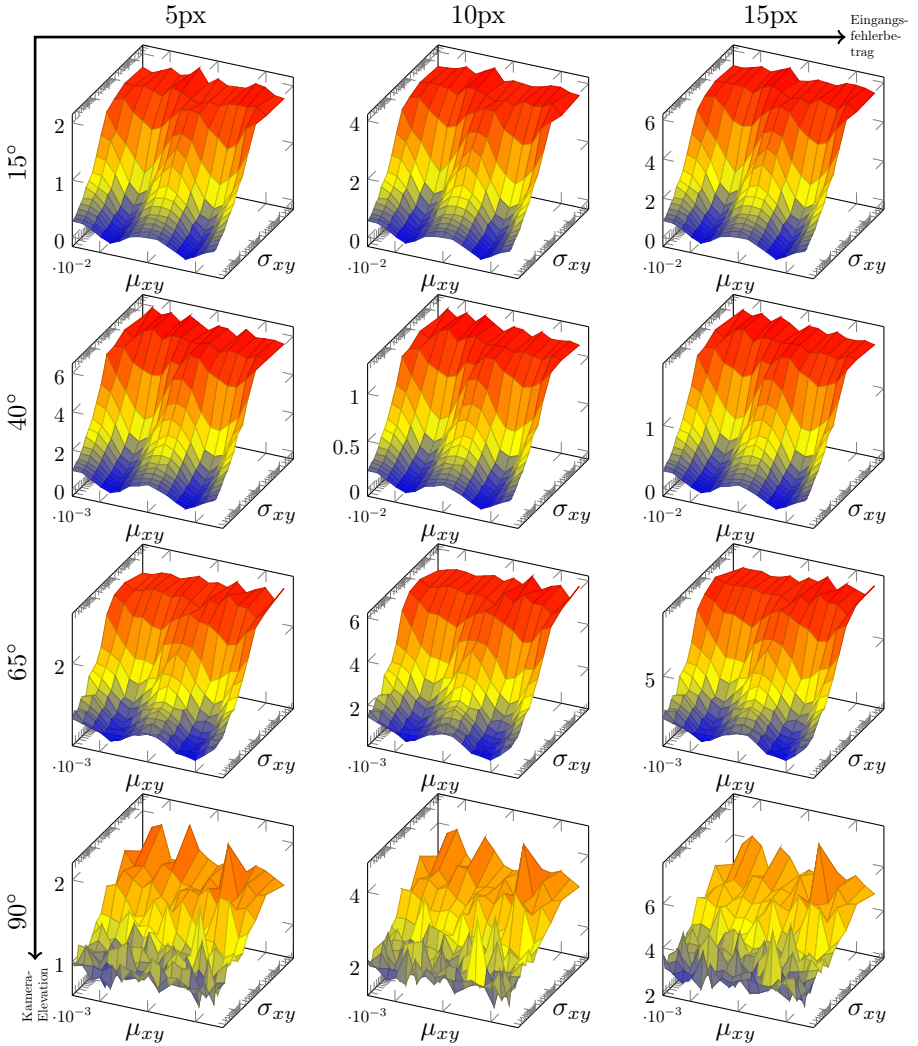


Abbildung 4.15: Vollabdeckung des Eingangsfehlerverteilungs-Parameterraumes stratifiziert nach Kamera-Elevation und Eingangsfehlerbetrag für die Ellipsenmittelpunkte

**Ellipsen-Halbachsenstreckung: „Fall  $ab$ “** — Die  $\mu_{ab}$ - $\sigma_{ab}$ -Vollabdeckung der Ellipsen-Halbachsenstörung liefert die Fehlerflächen in Abb. 4.16. Es zeigt sich, dass die in Abb. 4.12 für  $\sigma_{ab} = 0$  angepasste Modellfunktion

$$\sqrt{p_{\mu_{ab}} \left| \cos \left( \mu_{ab} + \frac{\pi}{4} \right) \right|} \quad (4.32)$$

tatsächlich eine Extremum an Varianz (bzgl.  $\mu_{ab}$ ) über den gesamten  $\sigma_{ab}$ -Bereich darstellt. Wird  $\sigma_{ab}$  groß (die Verteilung nähert sich der Gleichverteilung an), werden diese Kurven flacher, da die Extrema der Punktverteilung sich herauszumitteln beginnen. Die in Abb. 4.13 für  $\mu_{ab} = 0$  gemachte Beobachtung, dass der Verlauf in Richtung  $\sigma_{ab}$  konstant ist (mit Wert  $p_{\sigma_{ab}}$ ), hält größtenteils auch für die übrigen  $\mu_{ab}$ .

Die konstante Fortsetzung des Modells (4.32) über alle  $\sigma_{ab}$  unterschätzt aus dem selben Grund für größere  $\sigma_{ab}$  den Fehler in Umgebungen von  $\mu_{ab} \in \{-\frac{3}{4}\pi, \frac{\pi}{4}\}$ .

**Fazit** — Die Analyse der Störungen der separaten Ellipsenparameter  $xy$ ,  $ab$  und  $\theta$  ist damit abgeschlossen. Als wichtigste Erkenntnisse sind zu notieren:

1. Die dominanten Faktoren für die Vorhersage der Ausgangsfehler sind der Absolutbetrag der Eingangsfehler und die Kamera-Elevation gegenüber der beobachteten Oberfläche.
2. Für separat gestörte Variablen  $xy$ ,  $ab$  und  $\theta$  wachsen für alle Kamera-Elevationen die Ausgangsfehler linear über linear wachsenden Eingangsfehlerbetrag.
3. Die Ausgangsfehlermaxima liegen für die separaten Störvariablen nicht bei denselben Kamera-Elevationen. Bei Störung der Mittelpunkt und der Halbachsen wächst der Fehler mit in Richtung Nadir steigender Kamera-Elevation, bei Störung der Rotation fällt er. Das ist insofern günstig, dass sich diese Fehler dann nicht in voller Stärke ergänzen.
4. Für die Mittelpunktstörungen treten die Ausgangsfehlermaxima im Bereich der gleichverteilten Eingangsfehler ( $\sigma_{xy}$  groß) auf und steigen sehr schnell auf dieses Maximum, bei den Halbachsen sind sie recht konstant über  $\sigma_{ab}$  und überschreiten nirgends das Maximum von (4.32).

Aus den Punkten 1 und 2 muss man folgern, dass bei unbekannter Lage der beobachtenden Kamera und unbekannten (oder über das Bild variablen) tatsächlichen Eingangsfehler-Absolutbeträgen das Maximum der ermittelten Modelle angenommen werden muss.

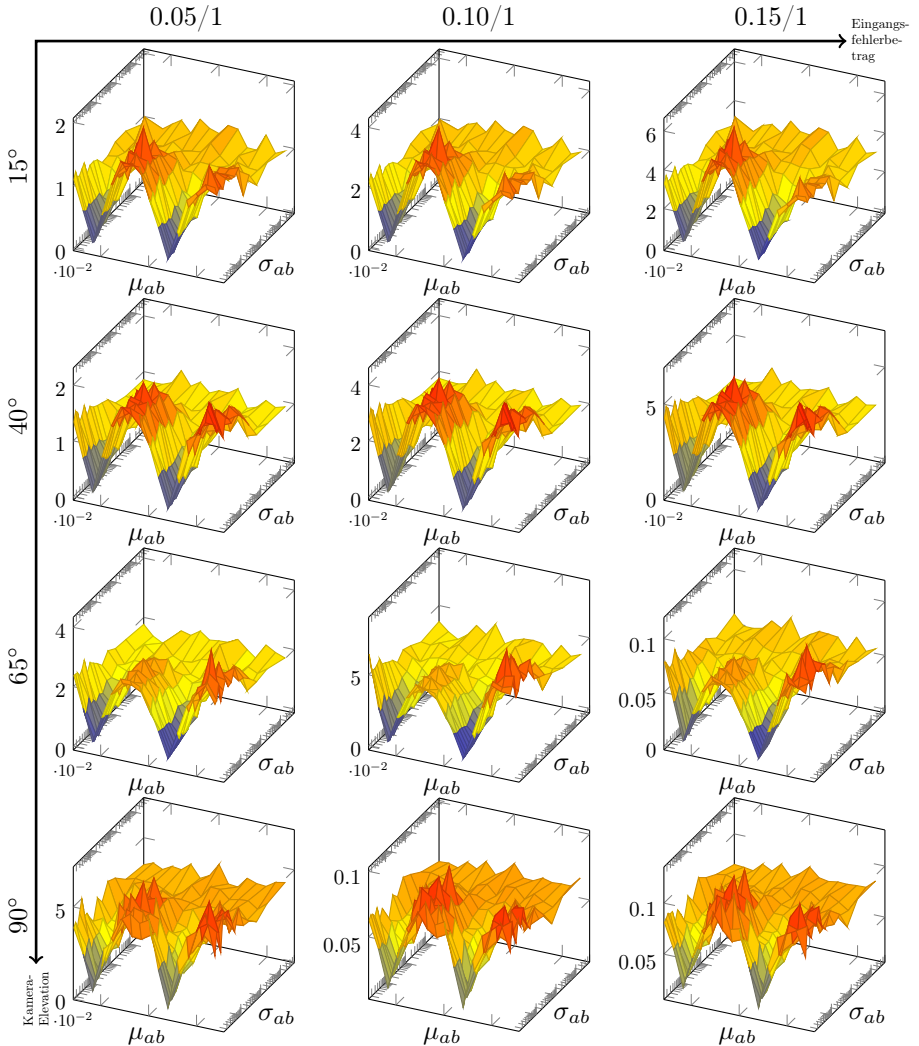


Abbildung 4.16: Vollabdeckung des Eingangsfehlerverteilungs-Parameterraumes stratifiziert nach Kamera-Elevation und Eingangsfehlerbetrag für die Ellipsen-Halbachsen

#### 4.3.4 Modellierte Schranke für gemischte Ausgangsfehler

---

Punkt 4 bedingt, dass für die Mittelpunktstörung anstatt des Modells (4.31) zur Beschränkung des Fehlers genausogut eine Grenzebene von konstant seinem Maximum verwendet werden kann, für die Halbachsen eine Grenzebene von konstant dem Maximum von (4.32). Das Modell des Rotationsstörungsfehlers ist ohnehin bereits konstant. Nur für sehr kleine  $\sigma_{xy}$  und schmale Bänder von  $\mu_{ab}$  wird der Fehler durch diese Vereinfachung signifikant überschätzt. Unterschätzt wird er praktisch nirgends.

Aus dem Punkt 3 folgt schließlich, dass das endgültige Ausgangsfehlermodell über das Kamera-Elevations-Argumentenmaximum der *Summe* der Einzelmodelle zu konstruieren ist.

#### 4.3.4 Modellierte Schranke für gemischte Ausgangsfehler

Es lassen sich mit gewisser Vorsicht aus den oben gemachten Untersuchungen notwendige Größen für eine Modellbildung über eine gemischte Störung aller Ellipsenparameter ableiten. Als formale Zusammenfassung der zuletzt gemachten Argumente folgt das Fehlermodell

$$\delta q_{\text{mod}} := p_{1,\sigma_{xy}}^{[\varphi_{\text{ref}}]} + p_{4,\sigma_{xy}}^{[\varphi_{\text{ref}}]} + \sqrt{p_{\mu_{ab}}^{[\varphi_{\text{ref}}]} + p_{\theta}^{[\varphi_{\text{ref}}]}}. \quad (4.33)$$

Die Größe  $\varphi_{\text{ref}}$  ist

$$\varphi_{\text{ref}} := \arg \max_{\varphi_n \in (0, \pi/2]} \left\{ p_{1,\sigma_{xy}}^{[\varphi_n]} + p_{4,\sigma_{xy}}^{[\varphi_n]} + \sqrt{p_{\mu_{ab}}^{[\varphi_n]} + p_{\theta}^{[\varphi_n]}} \right\}, \quad (4.34)$$

wobei  $\varphi_n$  der Kamera-Elevationswinkel aus (4.23) ist. Die ermittelten Parameter zu den Datensätzen der Abbildungen 4.10 bis 4.14, die für das Modell (4.33) benötigt werden, sind in Tabelle 4.3 aufgezählt. Die rechtsäußere Spalte enthält die Summe der Parameterwerte der vier links davon liegenden Spalten. Ihr Maximum ist die globale Schranke  $\delta q_{\text{mod}}$  aus (4.33).

Eine konstante Schranke für die Mediane der Ausgangsfehlerverteilungen wäre für die in den Tests dieses Kapitels verwendeten Eingangsfehlerbeträgen laut Tabelle 4.3 damit  $\delta q_{\text{mod}} \mathbf{2.59e-1}$  bei  $\varphi_{\text{ref}} = 15^\circ$ . An den Werten der letzten Spalte in den unteren drei Zeilen der Tabelle kann man allerdings auch schon ablesen, dass diese Schranke sehr konservativ ausfällt.

**Fazit** — Die Untersuchungen der vorangehenden Unterabschnitte und die hier gemachte Modellbildung lässt in die folgende Aussage bezüglich der Ausgangsfehlerverteilungen kondensieren:



#### 4.3.4 Modellierte Schranke für gemischte Ausgangsfehler

$\varphi_n [^\circ]$	$p_{1,\sigma_{xy}}^{[\varphi_n]}$	$p_{4,\sigma_{xy}}^{[\varphi_n]}$	$\sqrt{p_{\mu_{ab}}^{[\varphi_n]}}$	$p_{\theta}^{[\varphi_n]}$	$\Sigma$
15	2.65e-2	3.24e-2	4.03e-2	1.59e-1	2.59e-1
40	8.05e-3	1.02e-2	4.91e-2	4.94e-2	1.17e-1
65	3.60e-3	5.24e-3	7.34e-2	2.17e-2	1.04e-1
90	1.10e-3	4.14e-3	9.89e-2	5.73e-7	1.04e-1

Tabelle 4.3: Modellparameter für Ausgangsfehlermedian-Verteilungen über Kamera-Elevation bei maximalem Eingangsfehlerbetrag

- „ Im Fall maximaler gemischter Parameterstörungen
- a) der *Ellipsenmittelpunkte* in gleichverteilte Richtungen um fünfzehn Pixel,
  - b) der *Halbachsen* um 15 Prozent in jeweils einer Achse,
  - c) der *Rotation* der Ellipse um fünfzehn Grad in beliebig verteilte Richtungen in positivem oder inversem Umlaufsinne,
  - d) bei einer *Kamera-Elevation* von 15 Grad gegenüber der beobachteten Oberfläche,
- sollten mindestens 50 Prozent der relativen Fehler der verbindenden Strecken der rekonstruierten kamerafesten Kraterpositionen einen Wert von 26 Prozent nicht übersteigen.“

Dies ist eine konstante und sehr konservative Schranke, für größte Fehler mit ungünstigster Verteilung und bei flachstem Blickwinkel auf die Krater. In der Regel kann man erheblich kleinere Fehler erwarten, wie man in Tabelle 4.3 für die Kamera-Elevationen 40° und höher ablesen kann. Andererseits sind die Parameter aufgrund der unmodellierten Restfehler auch mit erheblichen Unsicherheiten behaftet und im Referenzszenario ist auch keinerlei Annahmen über die Kamera-Elevation zulässig. Von dem Standpunkt aus ist diese Vorsicht gerechtfertigt.

Die Modellvorhersage kann aber wieder teilweise überprüft werden, jedoch nur für eine sehr grobe Abtastung der Eingangsfehlerverteilungs-Parameterraumes von  $xy$ ,  $ab$  und  $\theta$ , da die kombinatorische Komplexität schnell überwältigend groß wird. Als Rechenbeispiel: Für die vollständig gemischte Verteilung der Eingangsfehler für jedes mögliche  $\mu_{xy}$  und  $\sigma_{xy}$  muss die volle Abtastung aller möglichen  $\mu_{ab}$  und  $\sigma_{ab}$  geprüft werden, und dafür wiederum al-

#### 4.3.4 Modellierte Schranke für gemischte Ausgangsfehler

---

le möglichen  $\mu_\theta$ . Es kommt daher bei einer Abtastung der Parameter  $\mu$  in acht Schritten und  $\sigma$  in zwei Schritten (die aus den in den Abschnitten 4.3.2 und 4.3.3 ermittelten Fehlerflächen folgenden minimalen sinnvollen Abtastungen) bereits eine Anzahl von  $8 \cdot 2 \cdot 8 \cdot 2 \cdot 8 = 2048$  Tests zusammen. Für jeden dieser Test müssen wiederum genug Stichproben aus der Eingangsfehlerverteilung gezogen werden, um diese hinreichend gut zu realisieren (Faktor  $\approx 30$ ), dann muss jeweils der ungestörte Fall und der maximal gestörte Fall berechnet werden (Faktor zwei), jeweils für alle Kamera-Elevationen (Faktor vier), und dann ist all dies immer noch für eine ausreichende Anzahl von Verschiedenen Testkraterkonfigurationen (Faktor  $\approx 30$ ) zu wiederholen. Diese Strategie würde also die Berechnung von 14.745.600 Splineinterpolationsproblemen mit dem dazugehörigen Auswertungsaufwand bedeuten.

Dünnt man die Abtastung der Eingangsfehlerverteilungen auf 16 Stichproben aus und die der Verteilung der Testkrater auf sechs Stichproben, reduziert sich dieser Umfang auf 1.572.864 Berechnungen, was zu bewältigen ist innerhalb einiger Tage Rechenzeit. Man erhält eine Ausgangsfehlerverteilung für jede getestete Kombination der Eingangsfehlerverteilungs-Parameter

$$\begin{aligned}\mu_{xy} &\in \{0, \pm \frac{\pi}{4}, \pm \frac{\pi}{2}, \pm \frac{3\pi}{4}, \pi\} \\ \sigma_{xy} &\in \{0, 50\} \\ \mu_{ab} &\in \{0, \pm \frac{\pi}{4}, \pm \frac{\pi}{2}, \pm \frac{3\pi}{4}, \pi\} \\ \sigma_{ab} &\in \{0, 50\} \\ \mu_\theta &\in \{0, \pm \frac{1}{4}, \pm \frac{1}{2}, \pm \frac{3}{4}, 1\}\end{aligned}\tag{4.35}$$

deren jeweiliger Median durch das Modell (4.33) vorhergesagt werden soll. Ergebnis ist eine Liste von 8192 Ausgangsfehlermedianen, verteilt zu je 2048 auf die vier Kamera-Elevationen.

Aufgrund der kleinen Stichproben ist das Ergebnis aber nur qualitativ zum Vergleich mit der Schranke (4.33) geeignet. In Abb. 4.17 ist dennoch sichtbar, dass die aus den wenigen Parametern der (teilweise nichtlinearen) Regressionsmodelle gewonnene Schranke tatsächlich eine gute Vorhersage des Verhaltens der Ausgangsfehler erlaubt, soweit im Rahmen der Aussagekraft der dünnen Eingangsfehlerabdeckung enthalten.

Dass die tatsächliche Verteilung der Mediane leicht unterhalb der modellierten Schranke streut, weist auf unmodellerte Wechselwirkungen der einzelnen Störvariablen hin, die den Gesamt-Ausgangsfehler abmildern. Überraschend ist es nicht, denn es gibt keinen offensichtlichen Grund, warum die Fehlerwirkung rein additiv sein sollte. Viel wichtiger ist, dass hiermit zumindest einen

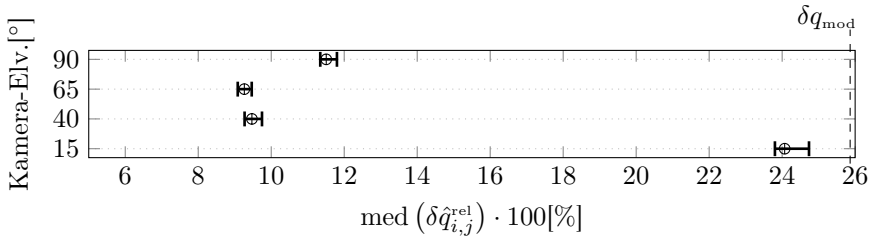


Abbildung 4.17: Modellierte Schranke und 97,5%-Quantile der realisierten Mediane der Ausgangsfehlerverteilungen für gemischte Eingangsfehlerverteilungen aller Störvariablen zu den vier getesteten Kamera-Elevationen mit 95%-Konfidenzintervallen

$\varphi_n$	$p_{1,\sigma_{xy}}^{[\varphi_n]}$	$p_{4,\sigma_{xy}}^{[\varphi_n]}$	$p_{\mu_{ab}}^{[\varphi_n]}$	$p_{\theta}^{[\varphi_n]}$
15	$m = +9.67\text{e-}4$	$m = +1.16\text{e-}3$	$m = +8.75\text{e-}3$	$m = +4.88\text{e-}1$
	$b = -4.59\text{e-}6$	$b = -3.40\text{e-}7$	$b = -3.68\text{e-}4$	$b = -4.39\text{e-}4$
40	$m = +3.05\text{e-}4$	$m = +3.72\text{e-}4$	$m = +1.35\text{e-}2$	$m = +1.48\text{e-}1$
	$b = +3.74\text{e-}6$	$b = +3.85\text{e-}6$	$b = -5.74\text{e-}4$	$b = -1.10\text{e-}4$
65	$m = +1.24\text{e-}4$	$m = +1.70\text{e-}4$	$m = +3.36\text{e-}2$	$m = +6.07\text{e-}2$
	$b = -1.37\text{e-}6$	$b = -1.33\text{e-}6$	$b = -1.29\text{e-}3$	$b = -7.18\text{e-}5$
90	$m = +7.56\text{e-}5$	$m = +1.54\text{e-}4$	$m = +4.47\text{e-}2$	$m = +4.54\text{e-}7$
	$b = -4.70\text{e-}5$	$b = -5.25\text{e-}5$	$b = -6.99\text{e-}4$	$b = +4.35\text{e-}7$

Tabelle 4.4: Ausgleichsgeraden über den Eingangsfehlerbetrag für die relevanten Parameter des Modells der Ausgangsfehlerverteilungs-Mediane

statistischen Hinweis gibt, dass diese Wechselwirkungen so gutartig bzw. klein sind, dass sie keine neuen inneren Maxima in der Ausgangsfehlerverteilung hervorrufen.

Möchte man (4.33) für größere als die hier betrachteten Eingangsfehlerbeträge extrapolieren, lassen sich die verwendeten Parameter anhand ihrer strengen Linearität über den Eingangsfehlerbetrag gut vorhersagen. Tabelle 4.4 zeigt die beobachteten Parameter der Geradengleichungen.

Es ergibt sich darüber hinaus als Seiteneffekt noch eine praktische Schranke für die Oberflächenrekonstruktion selbst. Die Iteration in den Gleichungen (3.104)–(3.105) in Unterabschnitt 3.7.9 muss beim Unterschreiten einer gewissen Schwelle der relativen Änderung von einer Iteration zur nächsten

abgebrochen werden. Mit den hier gefundenen Ergebnissen weiß man, dass man diese Schwelle schon bei  $10^{-3}$  ansetzen kann, denn das ist bereits eine Größenordnung genauer als der zu erwartende Fehler in den Daten.

## 4.4 Fehlerverhalten der Korrespondenzbildung

Durch Abdeckung des Raumes der Parameter von Verteilungen von Eingangsfehlern in den erkannten Ellipsen und Untersuchung des relativen Fehlers  $\delta q_{i,j}^{\text{rel}}$  (vgl. Gleichung (4.11)) in Abhängigkeit davon wurde der linke Zweig „ $\delta x$ “ in Abb. 4.3 auf Seite 142 abgedeckt. Für eine Bewertung der Performance des in dieser Arbeit vorgestellten Verfahrens ist aber der rechte Zweig, „ $\delta \pi$ “, von endgültigem Interesse.

Die Analyse im vorangegangenen Abschnitt ist hierzu direkte Voraussetzung, denn in der Hypothesenüberprüfung in Algorithmus 9 sind konstante Toleranzschwellen zu setzen, die sich genau auf Größen beziehen, die von den Tiefenfehlern der rekonstruierten Oberfläche abhängen. Wählt man diese beliebig, ist es unmöglich, die Fehler des Korrespondenzbildungsverfahrens aus Algorithmus 11 klar den nicht-optimalen Schwellen oder den Fehlern der Oberflächenrekonstruktion zuzuschreiben.

Per seiner Konstruktion ist der Zuordnungsalgorithmus letztendlich robust gegenüber schlecht gesetzten Schwellen, was falsche Positive angeht, da jede endgültige Hypothese erst nach Prüfung der Rückprojektionsresiduen zugelassen wird. Falsche Negative sind aber durch zu pessimistisch gewählte Schwellen durchaus möglich, es können also insbesondere korrekte Hypothesen fälschlich verworfen werden. Verwirft der Algorithmus dadurch alle korrekten Hypothesen oder zumindest so viele, dass seine Laufzeit über alle Maßen wächst, muss das als Versagen des Verfahrens gewertet werden.

Um systematisch untersuchen zu können, an welchen Punkten des Algorithmus das Korrespondenzbildungsverfahrens aus welchen Gründen versagt, werden in Tabelle 4.5 eine Reihe von Fehlerarten kategorisiert. Dabei ist die entscheidende Schwelle der Algorithmen benannt und inwieweit ihr Abweichen von den optimalen Werten das Versagen hervorruft.

Tabelle 4.5: Fehlerkategorien der Hypothesenprüfung in Alg. 11 des Korrespondenzbildungsverfahrens

Kategorie	Art	Fehler und ursächliche Abweichung
Falsche Positive	1	<code>hypo_rel_radii_are_consistent()</code> verwirft falsche Hypothese nicht: $tol_{\text{relrad}}$ zu groß (Gl. (3.116))
	2	<code>hypo_graph_angles_are_consistent()</code> verwirft falsche Hypothese nicht: $tol_{\text{relang}}$ zu groß
	3	Falsche Hypothese besteht Überprüfung der Reprojektionsresiduen: $tol_{\text{repres}}$ zu groß (Alg. 8)
Falsche Negative	4	<code>hypo_rel_radii_are_consistent()</code> verwirft korrekte Hypothese: $tol_{\text{relrad}}$ zu klein (Gl. (3.116))
	5	<code>hypo_graph_angles_are_consistent()</code> verwirft korrekte Hypothese: $tol_{\text{relang}}$ zu klein
	6	Korrekte Hypothese besteht Überprüfung der Reprojektionsresiduen nicht: $tol_{\text{repres}}$ zu klein (Alg. 8)
	7	Korrekte initiale 3-Hypothese hat zu großen MAHALANOBIS-Abstand zum Referenztripel
	8	Korrekte Hypothese wird wegen Erreichen einer harten Laufzeitgrenze nicht mehr gefunden

Nur, wenn in einem Lauf des Verfahrens die Fehlerarten 1 bis 3 der Kategorie „Falsche Positive“ zusammen bzw. hintereinander auftreten, wird das Verfahren eine irrtümliche Korrespondenztabelle  $\pi$  liefern. Diese Möglichkeit muss durch Wahl der Toleranzschwellen praktisch ausgeschlossen werden, da die Folge wäre, dass das Navigationssystem eine fehlerhafte Position und Lage liefert.

Der umgekehrte Fall, wenn die Fehler 4 bis 6 der Kategorie „Falsche Negative“ in Reihe auftreten, wird eine an sich korrekte Korrespondenz nicht als solche erkannt. Auch kann im Fall 8 eine an sich korrekte Hypothese gar nicht erst geprüft werden, wenn der Algorithmus vorher seine Laufzeit mit der Prüfung falscher Hypothesen verbracht hat. Die Struktur des Algorithmus 11 ist nämlich so angelegt, dass in jedem Lauf eine große Anzahl korrekter Hypothesen getestet werden, von denen mit großer Wahrscheinlichkeit trotz

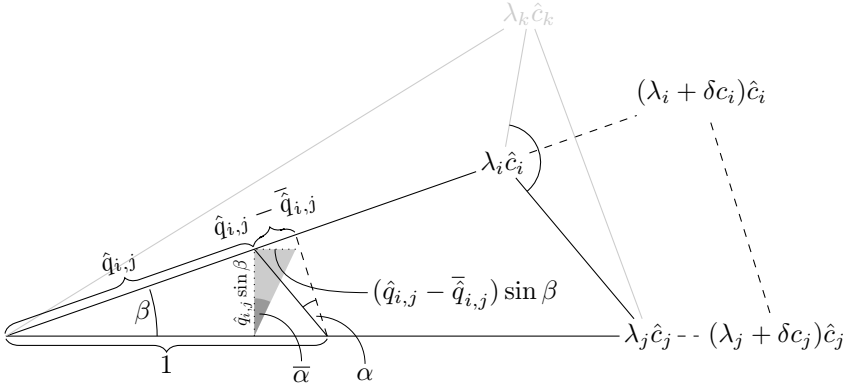


Abbildung 4.18: Approximation der Störung des Graphen-Innenwinkels durch Fehler in der Tiefenrekonstruktion an einem Hilfsdreieck und Zusammenhang mit ihren relativen Fehlern

Fehlern eine die Plausibilitätstests erfolgreich bestehen kann. Jedoch kostet die Überprüfung von mehr Hypothesen natürlich auch mehr Zeit und kann so die Laufzeit des Verfahrens über das zulässige Maß hinaus steigern. Auch deswegen dürfen die Toleranzschwellen nicht zu eng gewählt werden.

#### 4.4.1 Toleranzschwellen der Korrespondenzbildung

Mit der Analyse des Ausgangsfehlers der Oberflächenrekonstruktion im Abschnitt 4.3 und den dort gewonnenen Erkenntnissen über die real auftretenden Fehler in der rekonstruierten Oberfläche können jetzt konkrete Werte für die in Tabelle 4.5 genannten Schranken abgeleitet werden.

**a)  $tol_{relang}$ :** Zu den kamerafesten rekonstruierten Koordinaten  $\lambda_i \hat{c}_i$ ,  $\lambda_j \hat{c}_j$  und  $\lambda_k \hat{c}_k$  (mit jeweils  $\lambda_{\{i|j|k\}} := \hat{r}(\xi_{\{i|j|k\}})$ ) von drei beobachteten Kratern ist von Interesse, wie weit der Winkel  $\angle(\lambda_j \hat{c}_j - \lambda_i \hat{c}_i, \lambda_k \hat{c}_k - \lambda_i \hat{c}_i)$  gestört wird, wenn sich die Radien  $\lambda_i$ ,  $\lambda_j$  und  $\lambda_k$  um die Fehler der Oberflächenrekonstruktion  $\delta c_i$ ,  $\delta c_j$  und  $\delta c_k$  ändern. Als erste Vereinfachung lässt sich mit Sicherheit sagen, dass der gesuchte Fehlerwinkel kleiner ist als die Summe der Winkel

$$\begin{aligned} & \angle(\lambda_j \hat{c}_j - \lambda_i \hat{c}_i, (\lambda_j + \delta c_j) \hat{c}_j - (\lambda_i + \delta c_i) \hat{c}_i) \\ & + \angle(\lambda_k \hat{c}_k - \lambda_i \hat{c}_i, (\lambda_k + \delta c_k) \hat{c}_k - (\lambda_i + \delta c_i) \hat{c}_i) \end{aligned}$$

zwischen den gestörten und ungestörten Verbindungsstrecken.

Abb. 4.18 zeigt die Konfiguration für drei Beispielpunkte und den Zusammenhang mit den Größen des relativen Fehlermaßes (4.11) auf Seite 137. Insbesondere findet sich der Quotient in den Katheten des eingezeichneten Hilfsdreiecks wieder. Der Störwinkel  $\alpha$  zwischen den Strecken  $(\lambda_j + \delta c_j)\hat{c}_j - (\lambda_i + \delta c_i)\hat{c}_i$  und  $\lambda_j\hat{c}_j - \lambda_i\hat{c}_i$  wird dann approximiert durch den Winkel  $\bar{\alpha}$ , der mit den Quotienten  $q_{i,j}$  im Zusammenhang

$$\tan \bar{\alpha} = \frac{\sin \beta(\hat{q}_{i,j} - \bar{\hat{q}}_{i,j})}{\sin \beta \hat{q}_{i,j}} = \frac{\hat{q}_{i,j} - \bar{\hat{q}}_{i,j}}{\hat{q}_{i,j}} = \delta \hat{q}_{i,j}^{\text{rel}} \quad (4.36)$$

steht. Mit der Schranke für die Ausgangsfehlermediane (4.33) auf Seite 170, und da der Arkustangens monoton ist, gilt offensichtlich

$$\bar{\alpha} \leq \tan^{-1}(\delta q_{\text{mod}}) \quad (4.37)$$

für mindestens die Hälfte der Winkelfehler. Für die Paarung  $i$ - $k$  gilt dieselbe Schranke, es folgt also für den gesuchten Gesamtfehlerwinkel in Summe

$$\text{tol}_{\text{relang}} \approx 2 \tan^{-1}(\delta q_{\text{mod}}). \quad (4.38)$$

Für die in Unterabschnitt 4.3.4 benutzten Eingangsfehlerbeträge und die dadurch ermittelte Medianschranke wäre damit  $\text{tol}_{\text{relang}} \approx 30^\circ$ . An der Abb. 4.18 erkennt man, dass diese Schwelle geometrisch nur in Fällen ausgeschöpft wird, in denen  $\lambda_i \approx \lambda_j$  ist. In allen anderen Fällen ist  $\beta$  selbst sehr klein und diese Schwelle wird die tatsächlichen Fehler massiv überschätzen. Auch die Approximation von  $\alpha$  durch  $\bar{\alpha}$  enthält in den meisten anderen als dem abgebildeten Fall großzügige Marge. In späteren Tests wird sich zeigen, dass die praktische Schwelle

$$\text{tol}_{\text{relang}} := 8^\circ \quad (4.39)$$

tatsächlich optimal bezüglich der Erfolgsrate des Verfahrens ist.

**b)  $\text{tol}_{\text{relrad}}$ :** Der Ausdruck auf der rechten Seite von (3.116) in Unterabschnitt 3.9.6 enthält die skalierten Einheitsradien der beobachteten Krater. Ausgehend von der Annahme, dass die Radien zunächst nicht mit Messfehlern behaftet sind, ist der Ausdruck für jeden  $i$ -ten und letzten beobachteten Krater in einer aktuell betrachteten Konstellation äquivalent zu

$$\frac{(\lambda_i + \delta c_i)\bar{\rho}_i / (\lambda_{N_h+1} + \delta c_{N_h+1})\bar{\rho}_{N_h+1} - \lambda_i\bar{\rho}_i / \lambda_{N_h+1}\bar{\rho}_{N_h+1}}{\lambda_i\bar{\rho}_i / \lambda_{N_h+1}\bar{\rho}_{N_h+1}}. \quad (4.40)$$


---

#### 4.4.2 Test unter Vollabdeckung des Eingangsfehlerlandes

---

Umgestellt und mit der Notation  $\hat{r}(\xi_i) = \lambda_i$  und  $\bar{\hat{r}}(\xi_i) = \lambda_i + \delta c_i$  aus Unterabschnitt 4.2.1 wird dies zu

$$\left( \frac{\bar{\hat{r}}(\xi_i)}{\bar{\hat{r}}(\xi_{N_h+1})} \cdot \frac{\bar{\rho}_i}{\bar{\rho}_{N_h+1}} - \frac{\hat{r}(\xi_i)}{\hat{r}(\xi_{N_h+1})} \cdot \frac{\bar{\rho}_i}{\bar{\rho}_{N_h+1}} \right) \cdot \left( \frac{\hat{r}(\xi_i)}{\hat{r}(\xi_{N_h+1})} \cdot \frac{\bar{\rho}_i}{\bar{\rho}_{N_h+1}} \right)^{-1}. \quad (4.41)$$

Das wiederum ist nach Heraus Kürzen der Einheitsradien gerade der Ausdruck  $\delta \hat{q}_{i, N_h+1}^{\text{rel}}$  und die gesuchte Schwelle ist daher anzusetzen als

$$\text{tol}_{\text{relrad}} \approx \delta q_{\text{mod}}, \quad (4.42)$$

um 50% wahre Positive zu garantieren. Der konkret aus den oben gemachten Abdeckungstests abgeleitete Wert für diese Schwelle läge daher bei ca. 26%. Es zeigt sich jedoch in praktischen Tests, dass es in Balance mit der Laufzeitanforderung statthaft ist, deutlich mehr falsche Negative als 50% in diesem Test in Kauf zu nehmen, da der damit einhergehende Ausschluss von mehr falschen Positiven und der resultierenden „Irrwege“ entlang des Korrespondenzbaumes das Gesamtverfahren beschleunigt. Als praktisch optimale Schwelle stellt sich

$$\text{tol}_{\text{relrad}} := 17.5\% \quad (4.43)$$

heraus. Für beide oben abgeleiteten theoretischen Schwellen ließe sich über die Parametervorhersage-Tabelle 4.4 und das Modell (4.33) auch für andere Eingangsfehlerbeträge ein Richtwert generieren.

**c)  $\text{tol}_{\text{repres}}$ :** Diese Schwelle hängt nicht von den Ausgangsfehlern der Oberflächenrekonstruktion ab, sondern nur von den Eingangsfehlern in den Ellipsenmittelpunkten und vom Verhalten der EP $n$ P-QCP-Algorithmuskombination. Da diese in dieser Arbeit im Sinne eines „Off-the-Shelf“-Mechanismus eingesetzt wird, wird dazu hier keine tiefe Fehleranalyse vorgenommen. Als praktische Schwelle wird

$$\text{tol}_{\text{repres}} := 16\text{px} \quad (4.44)$$

gewählt. Gerade bei ungünstigen Kombinationen von Störungen und minimalen Anzahlen von Punktkorrespondenzen kann eine an sich korrekte Pose berechnet werden, die aber trotzdem ein Reprojektionsresiduum über dieser Schwelle erzeugt. Das muss als falsches negatives Ergebnis hingenommen werden. Wichtiger ist wie eingangs erwähnt, dass möglichst wenig irrtümliche Posen hinreichend kleine Reprojektionsresiduen erzeugen, um als korrekt bewertet zu werden.



### 4.4.2 Test unter Vollabdeckung des Eingangsfehler- raumes

Die konkreten in den Gleichungen (4.39), (4.43) und (4.44) gewählten Toleranzschwellen und die Parameterabtastung aus Abschnitt 4.3.4, Gleichung (4.35), werden jetzt benutzt, um das Verhalten der Korrespondenzbildung selbst zu untersuchen. Während im Abschnitt 4.3 mit der Größe  $\delta\hat{q}_{i,j}^{\text{rel}}$  ein nützliches Fehlermaß hergeleitet werden konnte, muss man sich hier auf eine rein frequentistische Betrachtung der Erfolgsrate zurückziehen.

Das bedeutet, dass jeweils nur Erfolg und Misserfolg der Korrespondenzbildung bewertet werden können, denn die Anzahl der Korrespondenzen ist entweder Null oder immer genau die Schnittmenge der Mengen sichtbarer, erkannter und bekannter Krater und daher bedeutungslos mit Hinblick auf die Erfolgsbewertung. Diese (bildweise) Gesamterfolgsrate ist in Abb. 4.19 abgetragen. Die für die Versuche gewählte Zulassungsschwelle für die Tripeldistanz aus Def. 3.14 ist 0.75 (relevant für Fehlerart 7 aus Tabelle 4.5). Die Laufzeitbeschränkung war eine dreivierteil Sekunde (relevant für Fehlerart 8).

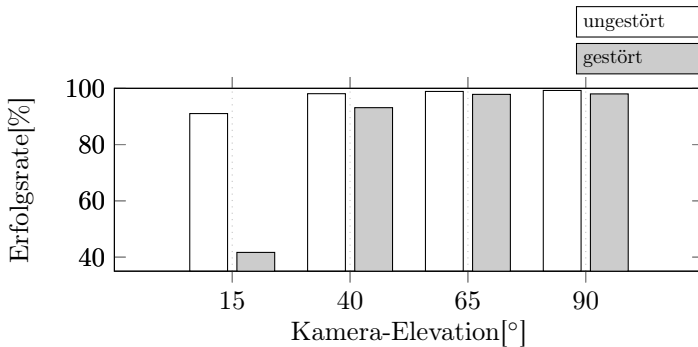


Abbildung 4.19: Erfolgsrate der Korrespondenzbildung im Vergleich zwischen ungestörtem Ergebnis und für gemischte Vollabdeckung der Eingangsfehlerverteilungsparameter mit maximaler Eingangsfehlermagnitude

Um detailliertere Statistiken über das Verhalten des Algorithmus 11 und die konkreten Gründe für die Misserfolge in Abb. 4.19 zu erhalten, werden seine Hilfsfunktionen und -Algorithmen so instrumentiert, dass sich gerade an den Schwellenvergleichen mit den oben hergeleiteten Toleranzen Erfolge und Misserfolge in die in der Tabelle 4.5 deklarierten Kategorien einordnen lassen.

Das auftreten dieser Fehlerkategorien ist (prozentual von allen Versuchen)

4.4.2 Test unter Vollabdeckung des Eingangsfehlerlandes

in Tabelle 4.6 erneut über die getesteten Kamera-Elevationen aufgetragen. Zu jeder Kamera-Elevation ist dort jeweils der Basiswert für den ungestörten Fall und darunter der Wert für den voll gestörten Fall abzulesen.

Kamera-elevation	gestört	Fehlerraten nach Art [%]							
		1	2	3	4	5	6	7	8
15°		1.06	1.07	1.62	35.41	66.34	32.75	73.21	0.16
	•	1.01	0.95	0.47	70.91	82.92	40.54	88.04	3.31
40°		1.02	0.96	1.54	40.39	24.16	27.55	56.75	0.01
	•	1.09	1.06	1.35	38.40	55.60	30.80	65.18	0.09
65°		1.10	1.09	4.11	24.61	8.01	21.81	19.63	0.00
	•	1.14	1.16	2.61	23.85	21.40	23.29	31.53	0.00
90°		1.32	1.21	4.80	3.34	3.45	0.12	1.18	0.00
	•	1.17	1.19	2.68	25.47	19.16	0.14	27.28	0.00

Tabelle 4.6: Prozentuale Fehlerraten der Zwischenstufen der Korrespondenzbildung in Kategorien nach Tab. 4.5

Folgende Beobachtungen und Erklärungen zu den tabellierten Werten:

**Fehler der Art 1)** Die Fehlerrate dieser Art bildet das Auftreten einer Art der falschen Positiven der erweiterten Hypothesenbewertung dar. Wie gewünscht kommen diese Fehler nur sehr spärlich vor. Im Detail lässt sich beobachten, dass bei kleinen Kamera-Elevationen die Durchlässigkeit der Radienvergleiche oft sogar kleiner ist, als für hohe Elevationen. Das macht daher Sinn, dass bei kleinen Kamera-Elevationen die relativen Fehler, wie im Abschnitt 4.3 beobachtet, größer werden und daher schneller die relevanten Toleranzschwellen überschreiten. Aus demselben Grund ist die Durchlässigkeit für den ungestörten Fall höher als für den gestörten. Im Sinne eines Klassifikators zeigen diese Werte, dass die verwendeten Tests eine sehr geringe Fehlalarm-Rate aufweisen.

**Fehler der Art 2)** Diese Art falscher Positiver entspricht in ihrem Verhalten denen der Art 1, da sie auf derselben Schwelle beruhen.

**Fehler der Art 3)** Dies ist der Anteil der irrtümlichen Hypothesen, die eine hinreichend plausible Kamerapose erzeugen, mit Hinblick auf die mit ihrer Hilfe rückprojizierten Katalogeinträge. In dieser Spalte stehen also genau die Anteile aller Versuche, in denen das Gesamtsystem eine falsche Kamerapose liefert. Wie erhofft, liegt diese Zahl gerade für die gestörten Daten um 2% oder teilweise deutlich darunter.

**Fehler der Art 4)** In dieser Spalte der Tabelle findet sich die Rate falscher Negativer der ersten Filterstufe für erweiterte Hypothesen wieder. Die relevante Toleranzschwelle in Gleichung (4.43) wurde niedriger angesetzt, als durch die theoretischen Überlegungen davor nötig, die Fehlerraten oberhalb von 50% bei 15° Kamera-Elevation reflektieren dies. Auffällig ist, dass auch die Eigenfehler der Oberflächenrekonstruktion (Approximationsfehler durch Polynomansatz, Restfehler durch Iterationsabbruch) bereits hohe Anteile der an sich korrekten Hypothesen nicht durchlassen. Man kann hierdurch indirekt darauf schließen, dass diese Abweichungen vielleicht nur eine Größenordnung unter den durch die gewollten Störungen der Ellipsen erzeugten Ausgangsfehler liegen.

**Fehler der Art 5)** Fehlerraten zur Art 5 zeigen, dass die zweite Filterstufe gerade in den gestörten Fällen ähnliche und nicht erheblich höhere oder niedrigere Raten von Fehlalarmen erzeugt wie die erste, ein wünschenswertes Verhalten.

**Fehler der Art 6)** Diese Fehlerraten geben wieder, dass i.A. ein bis zwei Fünftel aller korrekten finalen Hypothesen nicht die Überprüfung der aus ihnen abgeleiteten Kameraposen besteht, bis auf den nadirblickenden Fall, wo diese Fehlerart praktisch überhaupt nicht mehr auftritt. Dies ist einerseits wegen der künstlich eingebrachten Ellipsen-Mittelpunktstörungen zu erwarten, andererseits ist aber die Falsch-Negativ-Rate bei den ungestörten Systemen nicht erheblich kleiner. Das hier beobachtete Verhalten muss also Schwächen des EPnP-QCP-Verfahrens zugerechnet werden, da speziell im ungestörten Fall für korrekte Korrespondenzen an sich keine Fehler auftreten dürften. Man muss gleichzeitig aber bemerken, dass das hier konstruierte Testszenario mit Referenzebene ein pathologischer Fall aller solcher Verfahren zur Kameraposenbestimmung ist (der sog. „planare Fall“).

**Fehler der Art 7)** Die hohen Werte der Spalte 7 zeigen an, dass zu großen Anteilen korrekter initialer 3-Hypothesen die Sortierung nach dem Abstandsmaß aus Def. 3.14 diese Hypothesen aus dem Versuchsraum drängt. Experimentell wurde die Berücksichtigungsschwelle für dieses Abstandsmaß jedoch als bezüglich der Gesamtlaufzeit optimal bestimmt. Es scheint durch sehr enges Setzen dieser Schwelle ein Selektionsprozess stattzufinden, unter dem gerade diejenigen 3-Hypothesen passieren, deren Erweiterung auf größere Hypothesen eine hohe Erfolgsaussicht auf Passieren der nachfolgenden Filter haben (das Erstellen und Verwerfen vieler 3-Hypothesen ist günstiger als das Testen weniger Hypothesen von  $\geq 4$  Korrespondenzen).

**Fehler der Art 8)** Die letzte Spalte schließlich fasst die Gesamtausfallrate durch Überschreiten der festen Laufzeitgrenze (3/4 Sek.) zusammen. Man

erkennt, dass nur für kleinste Kamera-Elevationen überhaupt signifikante Ausfallraten auftreten. Insbesondere im gestörten 15°-Fall scheint die hohe Falsch-Negativ-Rate der Art 7 schließlich den optimalen Punkt zu passieren und es müssen ab und an zu viele Hypothesen getestet werden, als dass noch eine korrekte gefunden würde.

Zusammenfassend lässt sich sagen, dass die Kombination der Radien- und Winkelvergleiche zusammen mit der Prüfung der Rückprojektionsresiduen einen Klassifikator geringer Sensitivität und hoher Falsch-Negativ-Rate darstellt. Gleichzeitig, und erheblich wichtiger, besitzt er eine sehr geringe Fehlalarm-Rate. Da eine enormer Hypothesenraum zur Verfügung steht, wird selbst bei geringen Kamera-Elevationen noch einem großen Teil der Fälle eine korrekte Lösung ermittelt.

## 4.5 Zwischenfazit

Dieses Kapitel behandelte das Verhalten der Oberflächenrekonstruktion und der darauf aufbauenden Korrespondenzbildung mit Hinblick auf die Frage, was geschieht, wenn die Eingangsdaten (die Output-Ellipsen der Kratererkennungsalgorithmen) durch Bild- oder Algorithmenfehler gestört sind. Es wurden Überlegungen angestellt, welche Art von Störungen der Parameter in welchen Verteilungen modelliert werden müssten und ein Verfahren zur Abdeckung dieses Eingangsfehler-Parameterraumes definiert und implementiert.

Es stellte sich die Frage nach der Größe der Abweichungen der rekonstruierten Spline-Oberfläche in Abhängigkeit der so modellierten Eingangsfehler. Um eine Beschreibung dieser Abweichungen kompatibel mit den Algorithmen der Korrespondenzbildung zu machen, wurde das relative paarweise Fehlermaß  $\delta\hat{q}_{i,j}^{\text{rel}}$  definiert und es konnte demonstriert werden, dass unter Eingangsfehlern der (voher ermittelten) ungünstigsten Verteilung diese Ausgangsfehler nicht nur quantilweise beschränkt bleiben, sondern ihre Quantile auch nur linear wachsen mit dem Absolutbetrag der Eingangsstörung.

Aus diesen empirisch ermittelten statistischen Schranken wurden konkrete Werte für die bis dahin offenen Toleranzschwellen des Korrespondenzbildungs-Algorithmus abgeleitet. Unter Verwendung dieser Schranken wurde für die modellierten Eingangsfehler das gesamte Korrespondenzbildungsverfahren getestet und bezüglich Erfolgs- und Misserfolgsraten bewertet. Es stellte sich heraus, dass selbst für die als am ungünstigsten identifizierten sehr niedrigen Kamera-Elevationen die Erfolgsrate bildweise oberhalb von 40% bleibt und bei stumpferen Blickwinkeln sich bei 95% oder besser bewegt.

Da alle diese Tests unter Simulation der Outputs der Kratererkennung geschahen, verbleibt eine Restunsicherheit bezüglich der Möglichkeit, deren Ergebnisse und Fehler falsch eingeschätzt zu haben. Das folgende Kapitel wird daher einen Schritt weitergehen: Anstatt der Szenariogenerierung wird im geschlossenen Kreis das bildverarbeitende Kratererkennungsverfahren die Eingangsellipsen liefern, auf Basis von künstlich generierten Bildern realer Mondumgebung.



# Kapitel 5

## Tests des Gesamtverfahrens

### 5.1 Einleitung

In den Versuchen des Kapitels 4 wurde ermittelt, welche Verteilungen von künstlichen Störungen in den Kraterellipsen den größten Ausgangsfehler der Oberflächenrekonstruktion und damit die höchsten Versagensraten der Korrespondenzbildung erzeugen. Unter die dort modellierten Fehler fallen, wie schon in Abschnitt 4.2 aufgezählt, physische Makel der Kraterränder wie etwa Unterbrechungen, Mulden und Überschneidungen mit anderen Kratern, Abbildungsfehler der Kamera sowie algorithmische/approximative Fehler und Modellfehler in der Kraterrand-Rekonstruktion. In diesem Abschnitt soll ein weiterer Schritt in Richtung Realismus gegangen werden, indem die bildverarbeitenden Verfahren mit in die Tests aufgenommen werden.

Nach Ermittlung der Fehlerverteilungsparameter blieben als unabhängige und nicht weiter festgelegte Größen die Eingangsfehlerbeträge. Diese sind letztendlich auch im geschlossenen Kreis mit der bildverarbeitenden Kratererkennung schlicht nicht messbar, da die in dem Sinne „wahre“ Ellipse zu einem bestimmten Krater überhaupt nicht existiert: Jeder physische Kraterrand, der approximiert werden soll, ist schließlich mit Unsicherheiten bezüglich seiner Lokalisierung und durch die oben genannten physische Makel behaftet. Als weitere Faktoren kommen variable Beleuchtungsbedingungen hinzu, die den Kratererkennungsalgorithmus in noch unbekannter Weise beeinflussen.

Die im Abschnitt 4.4 ermittelten Erfolgsraten bewegten sich zunächst für die gestörten Daten (und hohe Kamera-Elevationen) oberhalb von 90%. Dabei blieben einige ganz wesentliche Faktoren aber unmodelliert, auch da

### 5.1.0 Einleitung

---

sie teilweise nicht sinnvoll parametrisierbar und dadurch als Zufallsgrößen darstellbar sind:

1) Auswirkungen des Zusammenspiels von Kameralage und Parametern der Beleuchtung. Steht z.B. die Sonne der Kamera genau „im Rücken“, so ist der Schatten in direkt in Blickrichtung liegenden Kratern unsichtbar und das Kratererkennungsverfahren versagt, obwohl an sich weder die Beleuchtungs-Elevation noch die Kamera-Elevation im kritischen Bereich liegen.

2) Einflüsse durch mangelhafte Abdeckung bzw. Dichte des Kataloges. Hierdurch wird dem Korrespondenzbildungsverfahren die Aufgabe gestellt, zwei mehr oder weniger disjunkte Mengen einander zuzuordnen, und in Abwesenheit tatsächlicher Korrespondenzen steigt die Wahrscheinlichkeit, dass eine irrtümliche Menge von Zuordnungen die Plausibilitätstests passiert. Nur, wenn überhaupt 4 bis 5 tatsächliche Korrespondenzen vorliegen und durch die rekonstruierte Oberfläche abgedeckt werden, ist eine erfolgreiche Zuordnung möglich.

3) Freie Stördaten, wie z.B. irrtümliche Erkennungen von Strukturen, die keine Krater sind. Hierunter fallen auch Bodenerhebungen oder Steine, die so zahlreich sind, dass sie die Schätzung der Beleuchtungsrichtung (vgl. Abschnitt 2.2) herumdrehen. Auch nicht-kreisförmige Krater sind in dieser Kategorie enthalten, da solche Krater erstens im mitgeführten Kraterkatalog per Definition nicht enthalten sein werden, und zweitens durch ihre Form die Rekonstruktion der Spline-Oberfläche erheblich stören können.

Die Differenz zwischen den in Abschnitt 4.4 bestimmten gestörten und ungestörten Erfolgsraten gibt einen Hinweis auf die Empfindlichkeit des Verfahrens gegenüber den dort untersuchten Störungen der Ellipsen. Die Differenz der Erfolgsraten der dort ermittelten gestörten Lösung und den Ergebnissen des folgenden Abschnitts 5.2 lässt sich dann in erster Näherung als die Wirkung der oben aufgezählten unmodellierten Einflüsse verstehen, wobei diese natürlich noch unquantifizierbare Annahmen bezüglich der Eingangsstörungsmagnitude enthalten und daher als untere Schranke verstanden werden müssen. Auf diese Weise sollen diese Wirkungen isoliert werden, um schließlich eine nützliche qualitative Aussage dazu abzuleiten, welche Umgebungs- und Missionsbedingungen im Allgemeinen welche Erfolgsaussichten versprechen. Dieser Teil der Untersuchungen ist Gegenstand des Abschnitts 5.2. An seinem Ende wird eine Aussage zur instantanen Erfolgsrate unter variierenden Kamera- und Beleuchtungsbedingungen, aber unter Konstanzhaltung des abgebildeten Geländes, stehen.

Ein weiterer wichtiger Test wird das Finden einer mittleren Akquisiti-



onsperiode sein. Damit ist die Zeitspanne gemeint, die man während eines Landeszenarios und seiner Orbitabschnitte<sup>1</sup> in Abhängigkeit von Flughöhe und Kameraparametern im Mittel warten muss, bis eine erfolgreiche Poseakquisition geschieht. Hiermit beschäftigt sich der Abschnitt 5.3. Der Vergleich der dabei anhand synthetischer Daten gewonnenen Erkenntnisse mit einem Versuch mit echter Kamera und einem physischen Geländemodell im Labor TRON ist dann Thema des Abschnitts 5.4.

## 5.2 Kamera- und Beleuchtungs-Elevation

In diesem Abschnitt werden die Bedingungen bezüglich der Kamera-Elevation so gut wie möglich abgedeckt, um den in 4 gefundenen starken Zusammenhang dieser mit der Erfolgsrate des Gesamtverfahrens zu bestätigen. Es werden dabei noch keine physikalisch sinnvollen Trajektorien der Kamera simuliert. Da das Kratererkennungsverfahren bildweise deterministisch ist und in Ermangelung einer Möglichkeit, beliebige realistische Mondgelände in gleichförmiger Weise zufällig zu generieren, ist die statistische Masse an Eingangsdaten für das Verfahren entsprechend kleiner (insbesondere fällt die in Abschnitt 4.3 ausgenutzte Variationsmöglichkeit für Kraterverteilung weg).

Zur Abmilderung dieser Einschränkungen wird daher für einen (möglichst repräsentativen) Bereich der Mondoberfläche eine regelmäßige Abdeckung des Raumes der möglichen Elevationen der beobachtenden Kamera erstellt, während gleichzeitig die bisher unbeachteten Beleuchtungsbedingungen der Szene in systematischer Weise variiert wird. Der angesprochene Mangel an Masse von Eingangsdaten wird dadurch so gut wie möglich ausgeglichen, dass zu jeder Kamera-Elevation eine regelmäßige Abtastung von Kamera-Azimuths und der genannten Beleuchtungsparameter hinzukommt. So wird zwar nach wie vor das selbe Gelände beobachtet, die Verteilung der Ellipsen im Bild und insbesondere das unbekannte Rauschen ihrer Parameter wird von Bild zu Bild neu abgetastet.

Das verwendete Geländemodell ist ein auf Basis von KAGUYA<sup>2</sup>-Stereometriedaten erstelltes DEM, es bildet einen Bereich von  $12 \times 24$  Grad (ca.  $400 \times 200$ km) des tatsächlichen Mondterrains ab mit einer lateralen Auflösung von ca. 100m/px. Ein Überblick hierüber ist in Abb. 5.1 gezeigt.

Dort ist auch der für diese Testserien tatsächlich abgebildete runde Aus-

---

<sup>1</sup>Vgl. m. d. Referenzzenario aus Abschnitt 1.1

<sup>2</sup>Japanische Mondmission mit mehreren optischen Instrumenten, für einen Überblick vgl. (Kato, Sasaki, Takizawa u. a., 2010).

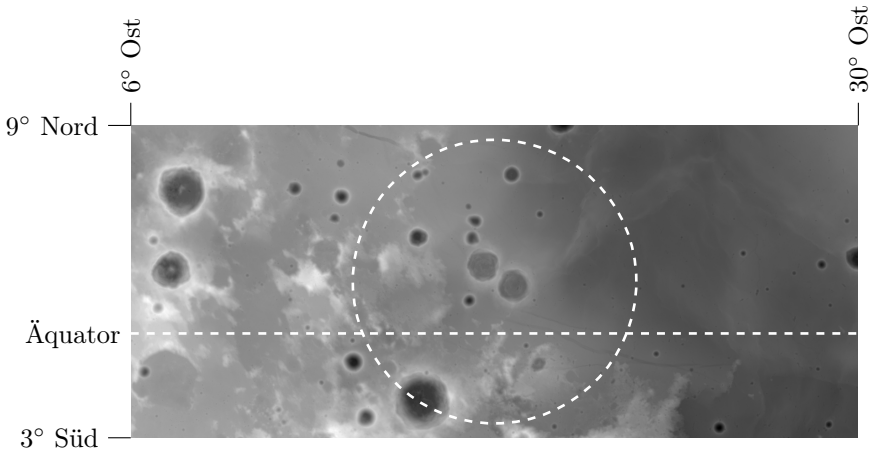


Abbildung 5.1: Geländeausschnitt des Mondes im Höhenmodell in simpler zylindrischer Projektion mit Kennzeichnung des ungefähren für die Tests verwendeten kreisförmigen Bereiches

schnitt eingezeichnet. Es wurde ein kreisförmiges Testgebiet gewählt, um für jede Kameraelelevation eine konstante Abdeckung des Bildbereiches mit tatsächlicher Information zu gewährleisten. Würde man das gesamte rechteckige Gelände abbilden, wären für gewisse Kamera-Azimuthwerte mehr (und anders verteilte) Ellipsen verfügbar, als für andere. Zusätzlich wird auf diese Weise der Bereich der tatsächlich über den Bildbereich realisierten Kamera-Elevationen begrenzt und so das Ergebnis relevant gehalten bezüglich der zugrundeliegenden Ziel-Elevation.

### 5.2.1 Der Posenraum von Kamera und Beleuchtung

In den Abdeckungstests des Abschnittes 4.3 wurde die Kamera-Elevation  $\varphi_n$  systematisch und regelmäßig von einem Minimalwert von  $15^\circ$  bis zum nadirblickenden Fall von  $90^\circ$  angehoben, immer mit Bezug auf die dort verwendete Referenzfläche  $S$ . Aufgrund der großen Skala des hier verwendeten Testgeländes ist die natürliche Mondkrümmung bereits relevant und zu berücksichtigen. Es wird daher  $\varphi_n$  von einem Wert von  $5^\circ$  gegenüber der zentralen Tangentialfläche an das Gelände bis zum Nadir variiert. Zur Vergrößerung der Masse an Eingangsdaten wird auch die Abtastdichte verfeinert auf Schritte

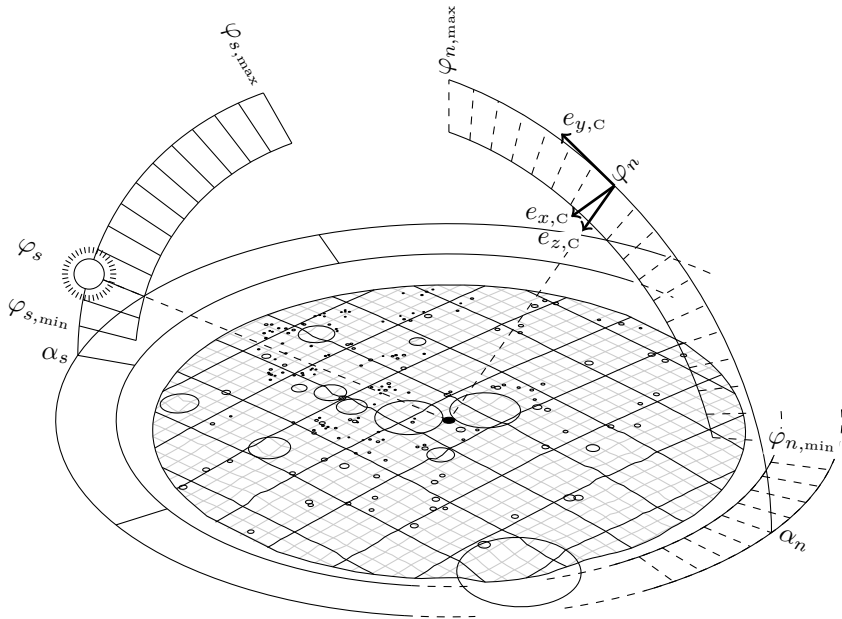


Abbildung 5.2: Kombinierte Variation von Kamera-Elevation, -Azimut, Sonnelevation und -Azimut für einen runden lunaren Geländeausschnitt von 200km Durchmesser

von  $5^\circ$  (18 Schritte). Der Kamera-Azimut  $\alpha_n$  wird seinerseits in 72 Schritte zu je  $5^\circ$  unterteilt.

Die Variation der Beleuchtung wird zu jedem solchen Azimut-Elevation-Paar der virtuellen Kamera selbst vollständig in Azimut  $\alpha_s$  und Elevation  $\varphi_s$  zwischen  $5^\circ$  und  $65^\circ$  variiert, jedoch mit 13 bzw. 5 Schritten wesentlich größer. Die hier beschriebenen Azimut- und Elevationsinkremente sind in Abb. 5.2 schematisch dargestellt. Insgesamt kommen auf diese Weise  $18 \cdot 72 \cdot 13 \cdot 8 = 134784$  einzigartige Kamera-Beleuchtungs-Konfigurationen zusammen, und damit genauso viele Testbilder.

Die optische Achse der virtuellen Kamera schneidet in jeder Konfiguration den zentralen Punkt des Geländemodells, die kamerafeste  $e_{y,C}$ - $e_{z,C}$ -Ebene steht stets orthogonal auf die Grundebene des Geländes. Anders als im Diagramm zu vermuten, ist die Beleuchtung nahezu parallel durch Simulation der Lichtquelle in quasi-unendlicher Entfernung.

Der Geländeausschnitt wurde so gewählt, dass er nicht bei allen Kamera-Azimutwerten vollkommen ähnlich aussieht. Auf diese Weise soll ein fälschlich überoptimistisches oder überpessimistisches Ergebnis verhindert werden durch versehentliche Anregung des Verfahrens mit optimalen oder auch extrem ungeeigneten Bilddaten. Ein zweitrangiger Nutzen hieraus ist, dass sich im Umkehrschluss ein Eindruck ergibt über die Arten von Gelände, die das Verfahren besser verarbeiten kann als andere.

## 5.2.2 Kraterkatalog

Der für die Korrespondenzbildung notwendige Kraterkatalog wurde mittels einer selbstgeschriebenen Software manuell am DEM aus Abb. 5.1 erstellt. Er enthält Krater aller Größenkategorien, die im Modell enthalten sind, jedoch wegen des manuellen Aufwandes jeweils nicht alle vorhandenen. Insgesamt sind ca. 150 Krater des verwendeten kreisförmigen Ausschnittes im Katalog abgebildet. Sie sind als Kreise auf dem Höhengitternetz der Abb. 5.2 eingezeichnet.

Der Katalog ist größer als das verwendete Gelände selbst, er enthält Krater, die entweder nicht im Blickfeld der Kamera liegen oder sogar außerhalb des kreisförmigen Ausschnittes. Das ist insofern realistisch, als im realen Anwendungsfall durch Mangel an Vorkenntnis über die genaue Position und Lage der Kamera ebenfalls ein erheblich größerer als der abgebildete Mondbereich in Betracht gezogen werden muss. Der Suchraum des Korrespondenzbildungsverfahrens wird dadurch auf realistischere Umfänge vergrößert. Damit wird damit der in der Aufzählung auf Seite 110 und unter Punkt 2) auf Seite 186 schon genannte mögliche Ausfallgrund relevant: Es werden im Regelfall viele Krater erkannt werden, die nicht im Katalog enthalten sind (und viele enthaltene Krater werden nicht erkannt werden). Ob das Verfahren gegen solche Fehler robust ist, ist von zentraler Bedeutung.

Der Maßstab 1:125.000 des DEM ist so groß gewählt, da dadurch die in den Daten unvermeidlich enthaltenen lateralen und Höhenfehler unterhalb der Auflösungsgrenze der virtuellen Kamera liegen. Somit ist auch die durch diese Datenfehler entstehende Ungenauigkeit des manuell erstellten Kraterkataloges entsprechend klein und muss nicht weiter als Fehlerquelle in Betracht gezogen werden.

### 5.2.3 Künstliche Bilddaten

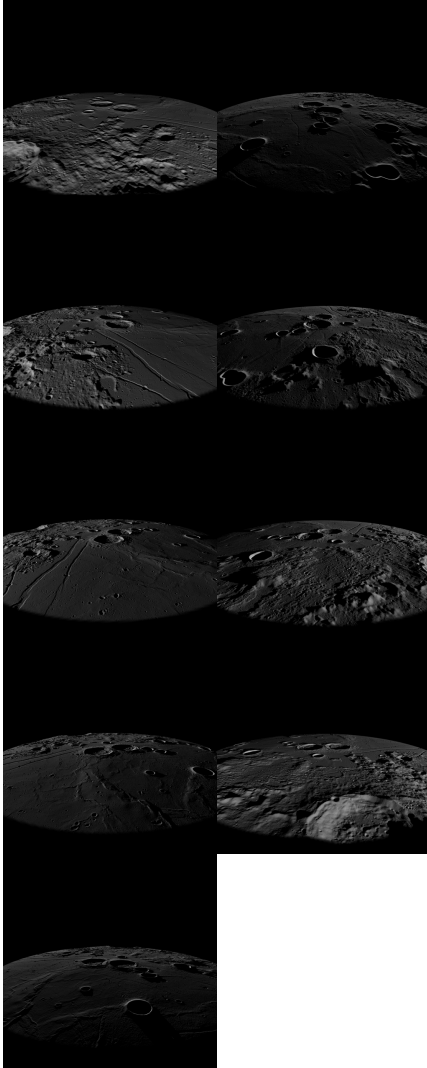
Die Testbilder des Geländes zu den oben genannten Kameraposen und Beleuchtungsbedingungen werden künstlich erzeugt, mittels Simulation der virtuellen Kamera in den oben beschriebenen Posen und Raytracings des mit einem Dreiecksnetz linear interpolierten Geländemodells. Die Software, die hierzu verwendet wird, wurde am DLR-Institut für Optische Sensorsysteme in Berlin erstellt (Paproth u. a., 2012). Mit ihrer Hilfe wurden auch die Bilddaten für die Softwaretests im Rahmen des DLR-internen Projektes ATON erstellt.

Beispiele für die resultierenden Bilder für einen Umlauf des Kamera-Azimuths bei konstanter Elevation und Beleuchtung sind in Abb. 5.3a gegeben, für ein Anheben der Kamera-Elevation bei konstantem Azimut in Abb. 5.3b. Die Abb. 5.4a zeigt den Effekt umlaufenden Beleuchtungsazimuths für konstante Beleuchtungselevation und konstante Kamerapose, die Abb. 5.4b demonstriert dann für die selbe Kamerapose aber konstanten Beleuchtungsazimut die Wirkung steigender Beleuchtungselevation auf die Bilder.

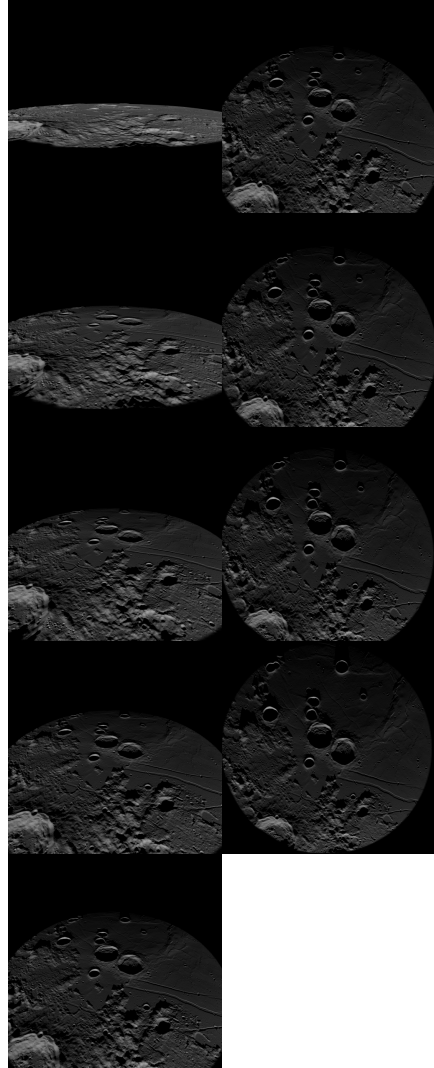
Es lässt sich in Abb. 5.3a erkennen, dass durch die Wahl eines kreisförmigen Geländeausschnittes die Ausfüllung des Bildrahmens mit tatsächlicher Information für jeweils konstanten Kamera-Azimut konstant bleibt. Bei Änderung der Kamera-Elevation ändert sich aber unvermeidlich auch der Anteil dieser Ausfüllung. Um zu verhindern, dass dadurch die Resultate bezüglich der Erfolgsrate verzerrt werden, werden die Ergebnisse der Testläufe zusätzlich zur Stratifizierung nach den Kamera- und Beleuchtungsparametern noch danach kategorisiert werden, ob die Anzahl der erkannten und auch im Katalog enthaltenen Krater überhaupt ausreichen würde, um eine Korrespondenztabelle zu ermitteln.

Weiterhin muss hier deutlich gemacht werden, dass die Erfolgsraten gerade für niedrige Kamera-Elevationen überkonservativ ausfallen werden, da ein sehr kleiner Teil des Bildes nur mit nutzbarer Information bedeckt ist, vgl. Abb. 5.3b. In einem realen Szenario ist das Bild immer bis zum Bildrand gefüllt (es sei denn, der Horizont ist im Blick). Dadurch würde erstens mehr Information für das Kratererkennungsverfahren verfügbar, und zweitens würde für die nahe dem Bildrand abgebildeten Punkte die effektive Kamera-Elevation (bezogen auf den Sichtstrahl durch die individuellen Bildpunkte) größer. Beides begünstigt eine höhere Zuordnungsrate, schwächt aber gerade die Abhängigkeit von der Kamera-Elevation ab.

Es wird eine Kamera mit einer virtuellen Brennweite von 778px simuliert, bei einer Bildbreite von 1024px entspricht das einem Field of View von  $66^\circ$ . Die Brennweite ist nicht Parameter der Tests, da von entzerrten Bildern aus-



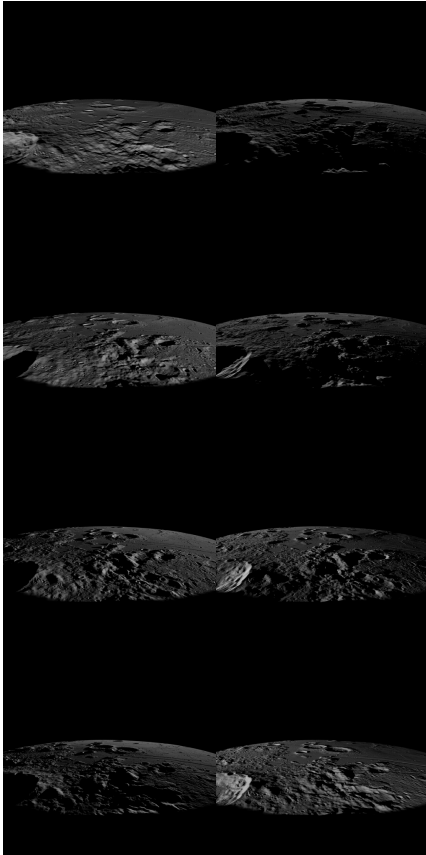
(a) Abtastung des Kamera-Azimuts in  $40^\circ$ -Schritten bei konstanter Elevation und Beleuchtung



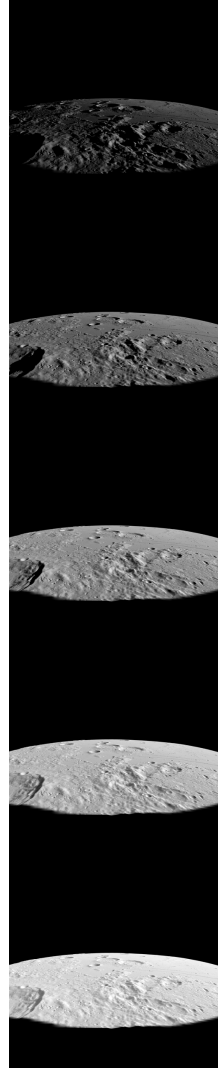
(b) Abtastung der Kamera-Elevation in  $10^\circ$ -Schritten bei konstantem Azimut und Beleuchtung

Abbildung 5.3: Auszug aus dem künstlichen Bildmaterial der systematischen Tests für Variation der Kamerapose

---



(a) Abtastung des Beleuchtungsazimuts in  $45^\circ$ -Schritten für eine konstante Kamerapose



(b) Abtastung der Beleuchtungselevation in  $10^\circ$ -Schritten für eine konstante Kamerapose

Abbildung 5.4: Auszug aus dem künstlichen Bildmaterial der systematischen Tests für Variation der Beleuchtungsbedingungen

gegangen wird und ihr Haupteinfluss daher in der effektiven Winkelauflösung der einzelnen Pixel besteht. Alle für die Algorithmen relevanten Werte werden jedoch in Pixeln gemessen und sind in diesem Sinne davon unabhängig. Erst bei extremen Brennweiten (sehr lang, entspricht orthografischen Kameras, oder sehr kurz, entspricht „Fischaugen“) würde der Einfluss auf die Oberflächenrekonstruktion erheblich wegen der Transformation in die virtuellen Bildkoordinaten.

#### 5.2.4 Ergebniskategorisierung

Es werden die selben Softwareparameter angewendet wie für die Störungsanalyse in Kapitel 4: Die Ausgleichsfläche wird über die acht größten erkannten Ellipsen rekonstruiert und es werden Korrespondenzen über mindestens fünf Krater gefordert, um eine valide Zuordnung anzunehmen. In allen folgenden Auswertungen werden die Ergebnisse der Testläufe über die Eingangsbilder unterteilt in die folgenden Kategorien:

- a) **Wahre Negative:** Eine gänzlich falsche Korrespondenztabelle in Fällen, in denen keine korrekte Zuordnung möglich war, da die Schnittmenge erkannter und bekannter Krater kleiner als fünf war.
- b) **Falsche Negative:** Eine falsche Korrespondenztabelle in Fällen, in denen grundsätzlich eine korrekte Zuordnung möglich gewesen wäre. Beinhaltet korrekte Korrespondenzen, die während der Plausibilitätstests der Korrespondenzbildung irrtümlich ausgeschlossen werden oder die die Schranken der Rückprojektionsbedingung überschreiten.
- c) **Wahre Positive:** Eine korrekte Korrespondenztabelle, die eine gute Poseschätzung der Kamera liefert.
- d) **Ambivalente Positive:** Eine teilweise falsche Korrespondenztabelle, die dennoch eine gute Schätzung der Kamerapose erlaubt.
- e) **Falsche Positive:** Eine gänzlich falsche Korrespondenztabelle, die eine grob falsche Kamerapose erzeugt, die aufgrund einer degenerierten Konstellation dennoch die Rückprojektionsbedingung erfüllt.

Als „Erfolg“ im Sinne der folgenden Ergebnisbesprechung werden Ergebnisse der Kategorien **c)** und **d)** gewertet. Die Erfolgsrate wird dann auf zweierlei Art berechnet: Zunächst als die Summe der Anzahl der Realisierung der Kategorien **c)** und **d)**, geteilt durch die Anzahl aller Versuche. Dann wird noch zum besseren Vergleich mit den Ergebnissen der Störungsanalyse diese Rate „korrigiert“ um die wahren Negativen der Kategorie **a)**: Die Erfolgsrate



berechnet sich in diesem Fall als die Summe der Anzahlen der Realisierungen der Kategorien **c)** und **d)**, geteilt durch die Anzahl aller Versuche abzüglich der Anzahl der Fälle, in denen keine korrekte Zuordnung möglich war aufgrund zu geringer Überschneidung von Erkennungen und Katalog. Diese Korrektur simuliert das Ergebnis für den hypothetischen Fall eines vollständigen Kraterkatalogs. Raten der ersten Art werden in den Balkendiagrammen weiß gefüllt. Raten der zweiten Art, grau.

### 5.2.5 Ergebnisse

Die Ergebnisse der Anwendung des Krateridentifikationsverfahrens auf die jeweils generierten Bilder zu den in Abschnitt 5.2.3 definierten Kameraposen und Beleuchtungsbedingungen wurde nach den oben genannten Kriterien eingeordnet. Die folgenden Betrachtungen stellen den Versuch dar, sowohl einen Gesamtüberblick über das Verhalten des Verfahrens zu schaffen, als auch besonders günstige oder ungünstige Parameter und Kameraposen zu identifizieren.

**Betrachtung nach Kameraposen:** Zunächst wird die Abhängigkeit der Erfolgsrate von der Kamera-Elevation überprüft. Abb. 5.5a zeigt die Erfolgsraten für die Variation der Kamera-Elevation von  $5^\circ$  bis  $90^\circ$ .

Im Vergleich zur Abb. 4.19 auf Seite 179 zeigt sich hier am Vergleichspunkt bei  $15^\circ$  Elevation ein ähnlicher Gesamtwert (korrigiert), aber ein erheblich niedrigerer Plateauwert bei Elevationen  $> 40^\circ$  insgesamt. Das ist auch nicht überraschend, denn im Kapitel 4 konnten zum einen nur die Verteilungen der Störungen identifiziert werden, die die größten Ausgangsfehler verursachen, während das obere Ende des Absolutbetrages der Störungen willkürlich festgelegt war, zum anderen konnten die „freien“<sup>3</sup> Fehler überhaupt nicht modelliert werden.

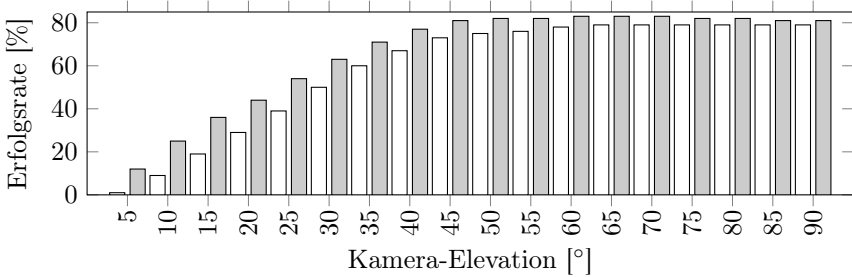
Das Profil der Erfolgsraten über die Elevation ist dafür ähnlich. Jeder Balken der Abbildung beinhaltet die Ergebnisse über einen kompletten Umlauf des Kamera-Azimuths und ist daher ein Mittelwert über die dahinter verborgenen Abhängigkeiten. Die Existenz von solchen Azimut-abhängigen Abweichungen wurde aufgrund der im letzten Absatz von Unterabschnitt 5.2.1 diskutierten Asymmetrie des Geländes vermutet. Abb. 5.5b zeigt tatsächlich, dass solche periodischen Abhängigkeiten existieren.

Die Azimutabhängigkeit ist im unkorrigierten Fall signifikant, die Unterschiede in der Erfolgsrate betragen gut 15% im Vergleich zwischen den Fällen, in denen von Süden bis Südosten auf das Gelände geschaut wird, zu

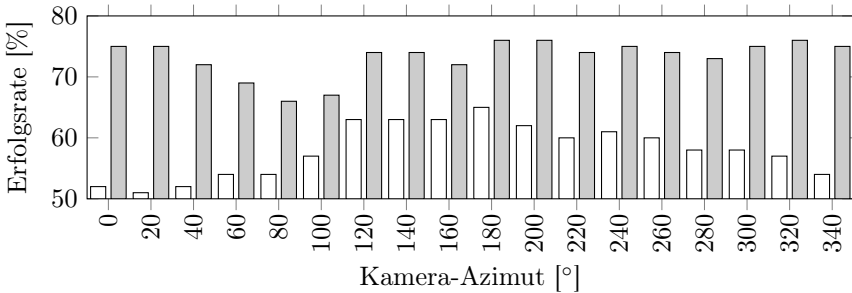
---

<sup>3</sup>Vgl. Punkt 3 in der Aufzählung auf Seite 186.

### 5.2.5 Ergebnisse



(a) Wahre und ambivalente Positive nach ansteigender Kamera-Elevation mit Bezug auf die mittlere Tangentialebene des Testgeländes



(b) Wahre und ambivalente Positive nach umlaufendem Kamera-Azimut gegenüber der lokalen Südrichtung

Abbildung 5.5: Erfolgsraten systematischer Tests stratifiziert nach den Poseparametern der Kamera (um die Anzahl der unmöglichen Fälle korrigierte Raten in Grau)

denen, in denen von Nord-Nordwest geschaut wird. Ein großer Teil dieser Abhängigkeit verschwindet aber nach Korrektur nach der Katalogdichte: Die grau gefüllten Balken zeigen nur direkt bei Westblick einen ähnlich großen Abfall der Erfolgsrate. Die in den weißen (unkorrigierten) Balken ablesbare beständige Schwäche bei Nordwest- bis Nordostblick liegt offenbar größtenteils daran, dass im bergigen südlichen Bereich des Testgeländes schlicht weniger Krater bekannt sind.

In Erweiterung dieser getrennten Betrachtungen der Kamera-Elevation und des Kamera-Azimuts lässt sich auch eine gemischte Betrachtung dieser Parameter machen. Der Vergleichbarkeit halber sind in den in Abb. 5.6 darge-

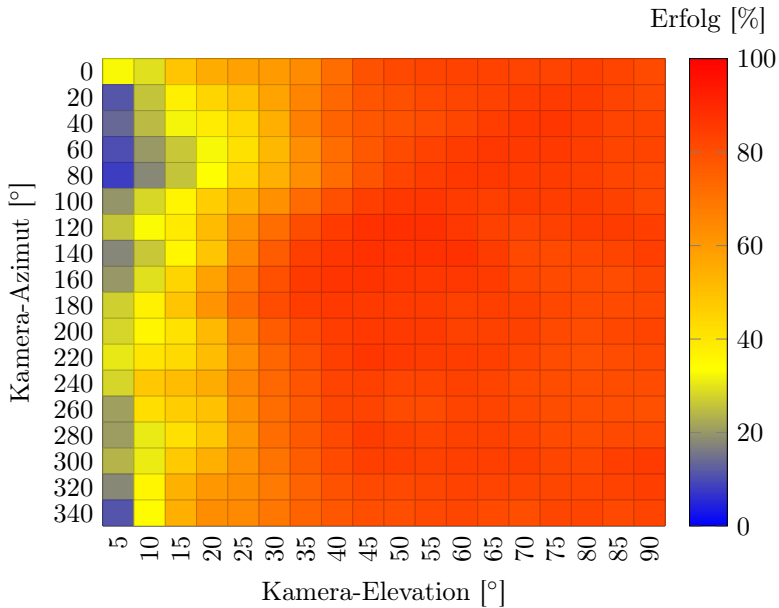
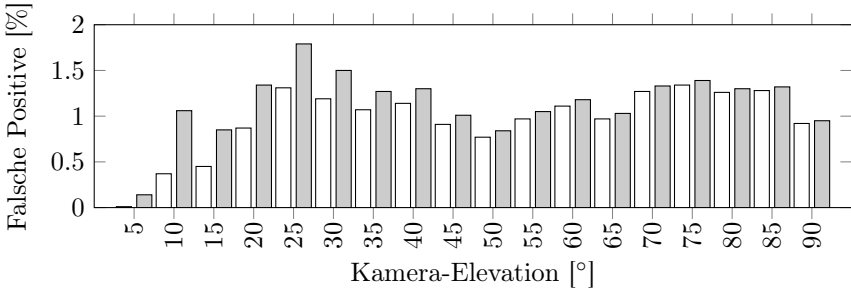


Abbildung 5.6: Erfolgsraten für volle Variation von Kamera-Azimuth und Kamera-Elevation

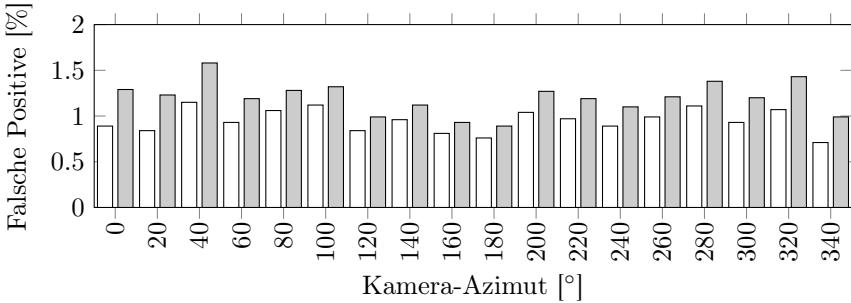
stellten Ergebnissen aber nur noch die um die Katalogabdeckung korrigierten Werte enthalten, die in den oberen Diagrammen als die grau gefüllten Balken abgetragen waren.

Die oben identifizierte Präferenz des Verfahrens für die südblickenden Fälle mit  $45^\circ$  Kamera-Elevation werden bestätigt. Man erkennt sogar, dass in diesem engen Szenario-Parameterbereich die Erfolgsraten gegen 90% gehen, obwohl weiterhin in jedem Datenpunkt des Diagramms jeweils der volle Bereich der Beleuchtungsparameter abgedeckt wurde. Ob sich diese Erfolgsraten durch Isolation optimaler Beleuchtungsparameter noch auf höhere Werte steigern lassen, wird später noch untersucht werden. Neben den Erfolgsraten ist die Falsch-Positiv-Rate von zentralem Interesse (Kategorie e) in der Aufzählung auf Seite 194). Betrachtet in Abhängigkeit von der Kamera-Elevation ist diese Größe in Abb. 5.7a abgetragen.

Als erfreulich soll zunächst herausgestellt werden, dass die Falsch-Positiv-Rate nirgends zwei Prozent überschreitet. Ob in den Daten zusätzlich



(a) Falsche Positive nach ansteigender Kamera-Elevation mit Bezug auf die mittlere Tangentialebene des Testgeländes



(b) Falsche Positive nach umlaufendem Kamera-Azimuth gegenüber der lokalen Südrichtung

Abbildung 5.7: Frequenz falscher Positiver stratifiziert nach den Poseparametern der Kamera (um die Anzahl der unmöglichen Fälle korrigierte Raten in Grau)

noch eine signifikante funktionale Abhängigkeit von der Kamera-Elevation vorhanden ist, ist aufgrund der kleinen Größenordnung dieser Fehlerraten und der großen beobachteten Varianz schwer zu fassen. Bei kleinen Kamera-Elevationen bis ca.  $25^\circ$  besteht immerhin eine gewisse Korrelation mit den Gesamterfolgswerten in dem Bereich. Diese Tatsache und der Umstand, dass die nach der Katalogabdeckung korrigierten Falsch-Positiv-Raten konsistent höher liegen als die unkorrigierten, weist darauf hin, dass immer noch einige an sich wahre Positive oder ambivalente Positive fälschlicherweise als falsche Positive eingeordnet wurden.

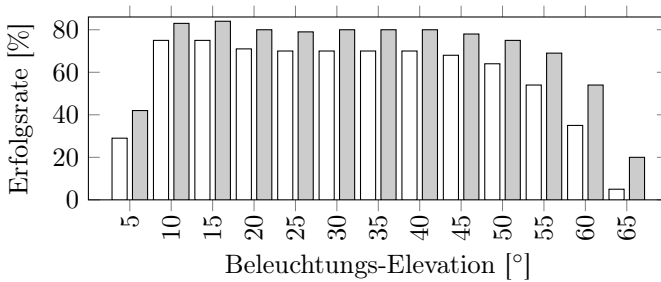
Die Abb. 5.7b zeigt dann die Falsch-Positiv-Raten, abgetragen über den Kamera-Azimet. Die Gesamtraten bewegen sich hier über den Parameterbereich um 1,5%. Es ist keine Korrelation mit den Erfolgsraten über den Kamera-Azimet in Abb. 5.5 erkennbar.

**Betrachtung nach Beleuchtungsparametern:** In den Diagrammen der Abb. 5.8 wird Wirkung der Variation von Elevation und Azimet der Lichtquelle gezeigt. Abb. 5.8a zeigt eine sehr markante Abhängigkeit der Gesamterfolgsrate des Verfahrens von der Beleuchtungs-Elevation. Wie zu erwarten war, ist die Erfolgsrate bei sehr niedrigen Einstrahlwinkeln klein, da hier die ausgeleuchteten Flächen innerhalb der Krater zu klein werden, um von den bildverarbeitenden Algorithmen berücksichtigt zu werden. Ebenso sinkt die Erfolgsrate für sehr hohe Einstrahlwinkel wieder ab, da hier zuerst harte Schlagschatten in den Kratern verschwinden und dann auch die aufgrund der einstrahlwinkelabhängigen Albedoparameter verbleibenden Kontrastunterschiede abnehmen. Die Erfolgsrate fällt schließlich bei ca.  $60^\circ$  Beleuchtungs-Elevation wieder unter den Wert der niedrigsten Elevation.

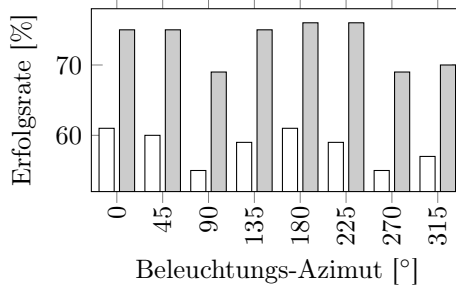
Das Erfolgsratenmaximum, das sich bei  $10^\circ$  bis  $15^\circ$  Beleuchtungs-Elevation beobachten lässt, approximiert das Optimum zwischen Ausleuchtung und Schärfe des Kontrastunterschiedes zwischen Hell- und Dunkel-Zonen der jeweiligen Krater und ihrer unmittelbaren Umgebung im Bild (vgl. Kap. 2, Abschn. 2.1). Auch, wenn sich dieses spezifische Optimum möglicherweise noch als abhängig von geografischen Variationen im Albedo herausstellt, ist dies ein wichtiger Hinweis auf eine Präferenz der Bildverarbeitung für die durch diese Beleuchtungs-Elevation erzeugten Bildeigenschaften und sollte bei der Missionsplanung unbedingt berücksichtigt werden.

Die Abhängigkeit vom Beleuchtungs-Azimet ist erheblich schwächer, die Variation in den Erfolgsraten beträgt über den gesamten Azimetbereich nur etwa 5%. Ein schwaches Minimum bei von Osten scheinender Sonne ist zu erkennen.

In der kombinierten Betrachtung von Elevation und Azimet der Beleuchtung in Abb. 5.9 lässt sich erkennen, dass die starke Ausfälle des Verfahrens auftreten, wenn die Lichtquelle sehr tief steht in einem Azimetbereich von  $270^\circ$  bis  $45^\circ$ , also von Süd-West bis Süd-Ost scheint. In diesen Fällen wird gerade die südliche Gebirgslandschaft angestrahlt und die effektive Beleuchtungs-Elevation in den östlichen und nördlichen Ebenen des Geländes sind nahe Null. Diese Flächen enthalten aber gerade den Großteil sowohl der detektierbaren als auch der bekannten Krater, die dann in Konsequenz der Beleuchtungssituation nicht erkannt werden. Auch bei den sehr hohen Beleuchtungs-Elevationen fällt



(a) Wahre und ambivalente Positive nach der Beleuchtungs-Elevation gegenüber der mittleren Tangentialebene



(b) Wahre und ambivalente Positive nach dem Beleuchtungs-Azimet gegenüber der lokalen Südrichtung

Abbildung 5.8: Erfolgsraten systematischer Tests stratifiziert nach den Beleuchtungsparametern der globalen Lichtquelle (um die Anzahl der unmöglichen Fälle korrigierte Raten in Grau)

die Erfolgsrate ab, weil wie erwartet keine Krater mehr detektiert werden.

Die Schwächen in der Erfolgsrate in den genannten Azimutbereichen der Beleuchtung verschwinden, sobald die Beleuchtungs-Elevation über  $10^\circ$  anwächst. Die oben angesprochenen Geländegebiete werden dann ausgeleuchtet und die dort befindlichen Krater werden detektiert. Die Performance des Verfahrens bleibt dann oberhalb dieser Beleuchtungs-Elevation sehr homogen, bis sie nach  $55^\circ$  wieder abfällt.

Abb. 5.10 fasst die Falsch-Positiv-Raten zusammen, stratifiziert nach den Parametern der Beleuchtung. Die Unterabbildung 5.10a zeigt ein markantes

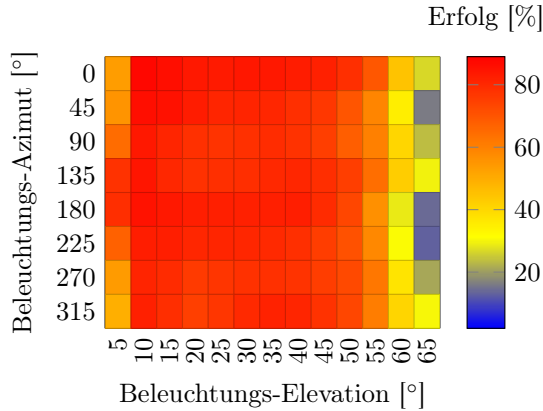


Abbildung 5.9: Erfolgsraten für volle Variation der Beleuchtungsparameter

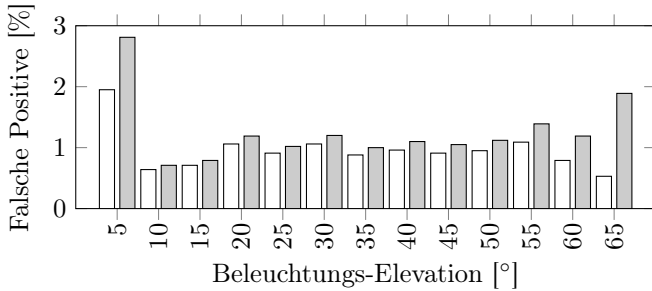
Profil, dass offenbar invers korreliert mit dem der Erfolgsraten aus Abb. 5.8a ist. Insbesondere die signifikanten Spitzen falscher Positiver bei 5° und 65° sind mit dem starken Abfall der Erfolgsraten bei diesen Beleuchtungs-Elevationen kongruent. Dass die Falsch-Positiv-Rate in den Fällen, in denen um die Katalogabdeckung korrigiert wurde, im Verhältnis hier höher liegt als bei den größeren Beleuchtungs-Elevationen, deutet wieder auf eine verbleibende Fehl kategorisierung von wahren (oder ambivalenten) Positiven als falschen Positiven hin.

Das Profil der Falsch-Positiv-Raten für umlaufenden Beleuchtungs-Azimat ist in Abb. 5.10b abgebildet. Es zeigt keine so signifikanten Spitzen oder Abfälle wie das der Beleuchtungs-Elevation und bewegt sich insgesamt im Bereich von ca. 1% bis 1.5%.

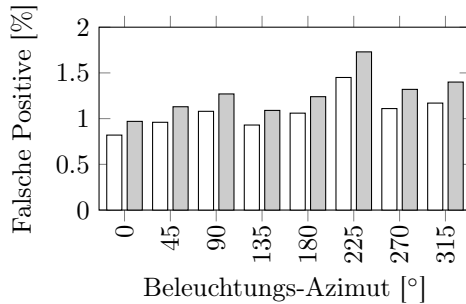
## 5.2.6 Auswertung und Fazit

Bei der Untersuchung der Ergebnisse wurden die Kamera-Elevation und die Beleuchtungs-Elevation wie erwartet als die Treiber der Dynamik in den Erfolgsraten bestätigt. Ihre kombinierte Betrachtung unter Vernachlässigung der jeweiligen Azimute ist daher am aussagekräftigsten. Abb. 5.11 zeigt die Erfolgsraten für die Kombination der beiden Parameter.

Es lässt sich deutlich ablesen, dass insbesondere die Kombination von



(a) Falsche Positive nach der Beleuchtungs-Elevation gegenüber der mittleren Tangentialebene



(b) Falsche Positive nach dem Beleuchtungs-Azimut gegenüber der lokalen Südrichtung

Abbildung 5.10: Falsch-Positiv-Raten stratifiziert nach den Parametern der Beleuchtung

niedrigen Kamera-Elevationen und hohen Beleuchtungs-Elevationen pathologisch ist für das Verfahren. Das erklärt sich daher, dass bei sehr spitzem Blick auf den Krater und sehr hoher Beleuchtung der noch schattierte bzw. beschienene Bereich in seinem Inneren unterhalb des kamerazugewandten Randes verschwindet. Die Schärfe des abgebildeten Kraterandes nimmt dann erheblich ab und der Kraterand-Deskriptor (vgl. Abschnitt 2.2) verliert an Darstellungskraft. Insgesamt lässt sich ein Bereich hoher ( $> 80\%$ ) Erfolgsrate abgrenzen, der den Bereich von  $10^\circ$  bis  $50^\circ$  Beleuchtungs-Elevation und  $35^\circ$  bis  $90^\circ$  Kamera-Elevation umfasst.

Hieraus lassen sich mit (wegen des konstant gehaltenen Testgeländes ge-



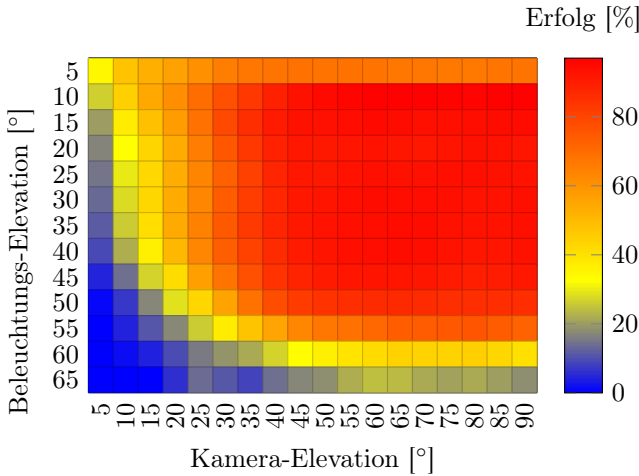


Abbildung 5.11: Erfolgsraten für die kombinierte Variation von Kamera- und Beleuchtungs-Elevation

botener Vorsicht) Schlüsse ziehen bezüglich optimaler Bedingungen für den Einsatz des in dieser Arbeit vorgestellten Verfahrens. Da die Rechenkapazität eines Navigationsrechners und das Zeitfenster, das zur Poseakquisition zur Verfügung steht, stark beschränkt sind, würde man im Rahmen einer hypothetischen Missionsplanung daher einen Bereich von Trajektorien- und Beleuchtungsparametern festlegen, der eine möglichst hohe Erfolgsaussicht (bildweise) bietet, um die zu erwartende Rechendauer bis zur erfolgreichen Poseakquisition innerhalb des zulässigen Rahmens zu belassen.

Einschränkend muss berücksichtigt werden, dass die Verteilung der Erfolgsraten in keinem hinreichenden Sinne unabhängig identisch verteilt ist über aufeinanderfolgende während einer stetigen Bewegung aufgenommene Bilder. Ist ein Bild aus irgendwelchen Gründen schlecht geeignet für die Anwendung dieses Verfahrens, muss man davon ausgehen (bei kleinen relativen Änderungen der Kamerapose), dass es auch das nächste sein wird. Außerdem ist nur ein kleiner Teil der tatsächlich realisierten Erfolgsrate als zufälliger Prozess modellierbar. Diese Tatsachen verhindern die Anwendung statistischer Modellbildung zur Ableitung von Konfidenzschwellen über die zu erwartenden Erfolgsraten aus den hier gesammelten Ergebnissen. Mit dieser Problematik beschäftigt sich der folgende Abschnitt 5.3.

5.3.0 Mittlere Akquisitionsperiode

Erwartete Erfolgsrate	Missionsparameter		
	Kamera- Elevation		Beleuchtungs- Elevation
niedrig ( $< 50\%$ )	$< 20^\circ$	oder	$< 10^\circ$ oder $> 55^\circ$
mittel ( $50\% - 75\%$ )	$20^\circ$ bis $35^\circ$	oder	$50^\circ$ bis $55^\circ$
hoch ( $> 75\%$ )	$> 35^\circ$	und	$10^\circ$ bis $50^\circ$

Tabelle 5.1: Erwartete bildweise Erfolgsraten in Abhängigkeit von den äußeren Szenarioparametern Kamera-Elevation und Beleuchtungs-Elevation

Als instantane Erfolgswahrscheinlichkeit betrachtet lässt sich aber eine grobe Eingruppierung in Erwartungsbereiche bezüglich der Erfolgsrate aus der oben gemachten Auswertung der Ergebnisse ableiten. Diese sind in Tabelle 5.1 aufgelistet.

Die in der Tabelle zusammengefassten Erkenntnisse sind das zentrale Resultat dieses Abschnittes. Die verbleibende Unsicherheit bezüglich der Anteile der in Kap. 4 modellierten Störungen und den in diesem Kapitel hinzugekommenen Fehlerquellen an den letztendlich realisierten Versagensraten verbleibt aber.

5.3 Mittlere Akquisitionsperiode

Es bleibt offen, wie sich die Erfolgsraten während einer realistischen Landetrajektorie verhalten würden. Hier tritt genau die oben besprochene Problematik in den Vordergrund, dass ein solches Szenario eine Abfolge von Bildern erzeugt, deren Eignung für eine erfolgreiche Krateridentifikation autokorreliert. Intuitiv ist das einleuchtend, denn ein Großteil der Erfolgswahrscheinlichkeit hängt vom abgebildeten Gelände ab. Um eine in einem stochastischen Sinne unabhängige Versuchsreihe zu erhalten, muss also streng genommen zwischen zwei Messungen das gesamte abgebildete Gelände ausgetauscht sein. Innerhalb welcher Zeiträume das passiert, hängt von den Parametern der Kamera und ihrer Trajektorie ab:

- Der Öffnungswinkel bestimmt die Sichtfeldgrenzen auf der Oberfläche.
- Die Kamera-Elevation bestimmt, wie weit das Sichtfeld der Kamera für

diesen Öffnungswinkel reicht, insbesondere, wie weit in Richtung des Horizontes die entfernte Sichtfeldgrenze liegt (bei „aufrechter“ Kameralage, wenn sich der Nadir im Bild unten befindet).

- Die Geschwindigkeit und Höhe der Kamera über dem Gelände beeinflussen die Geschwindigkeit des optischen Flusses des Bildes.

Da die Algorithmen des Krateridentifikationsverfahrens auch innerhalb der Periode eines kompletten Austausches des Bildinhalts durch perspektivische Effekte und das Herein- und Herauswandern von Kratern in das Bild und aus dem Bild Variation ihrer Eingangsdaten erfahren, finden zu jedem Bereich von großenteils sich auf diese Weise einander ähnelnden Bildern mehrere Messungen statt, die dann aber nicht voneinander unabhängig sind. Die Korrelation liegt hier in der Menge der über die Versuche übereinstimmenden detektierten Krater. Bei entsprechenden Öffnungswinkeln und Elevationen der Navigationskamera sind also möglicherweise nur sehr wenige oder überhaupt keine Messungen voneinander im oben beschriebenen Sinne unabhängig, z.B. wenn der Horizont im Bild liegt und der Landeplatz sich noch davor befindet.

Für die Missionsplanung muss dennoch ein Maß gefunden werden, in Abhängigkeit dessen sich benennen lässt, „wie lang“ es durchschnittlich dauern wird, bis das in dieser Arbeit entwickelte Verfahren eine gültige Initiallösung für den Zustandsschätzer liefert. Dieses durchschnittliche Intervall soll von hier ab als die *mittlere Akquisitionsperiode* bezeichnet sein. Sie wird gemessen in ihrer natürlichen Einheit „Bild“, denn erstens sind ohne Vorwissen über die Position, Lage oder Geschwindigkeit des Raumfahrzeugs alle metrischen Angaben bedeutungslos, und zweitens ist die benötigte Rechenzeit direkt von der Anzahl von Bildern abhängig, die bis zur erfolgreichen Akquisition verarbeitet werden müssen.

### 5.3.1 Bildaustauschperiode

Die oben genannten Kamera- und Trajektorienparameter Elevation, Flughöhe, Öffnungswinkel und Geschwindigkeit lassen sich betrachten als die unabhängigen Variablen einer Abbildung auf die mittlere Akquisitionsperiode als abhängige Größe. Für den Referenzfall Mond ist das eine gute Näherung. Denkbar ist aber auch, dass Eigenbewegung oder -Rotation eines abgebildeten Kleinhimmelskörpers Bildänderung verursacht. Um sich von spezifischen Trajektorien- und Umgebungsbedingungen zu lösen, ist es hilfreich, eine entdimensionalisierte Zwischengröße einzuführen:

**Definition 5.1 (Bildaustauschperiode)** Zu einer gegebenen zeitabhängigen Trajektorie einer Kamera bezeichne die Bildaustauschperiode das Trajektorienintervall eines vollständigen Austausches des abgebildeten Geländes durch die Relativbewegung der Kamera oder des abgebildeten Himmelskörpers.

Diese Definition macht es möglich, bei der Ermittlung der mittleren Akquisitionsperiode die Bildüberlappung und damit den Abhängigkeitsgrad der Einzelmessungen zu quantifizieren und als weitere unabhängige Variable neben der Kraterkatalogdichte in die Untersuchungen aufzunehmen. Die Bildaustauschperiode ist unabhängig von der konkreten Trajektorie, jedes Szenario mit einer Trajektorie lässt sich aber in Bildaustauschperioden umrechnen durch Berücksichtigung der oben genannten Parameter der Kamera und ihrer Bewegung und die in diesem Abschnitt gewonnenen Erkenntnisse sind gültig. Wie direkt diese Umrechnung geschieht, hängt letztendlich von der Geometrie des Geländes ab und ihrer Interaktion mit dem Kamerasichtfeld während der Trajektorie.

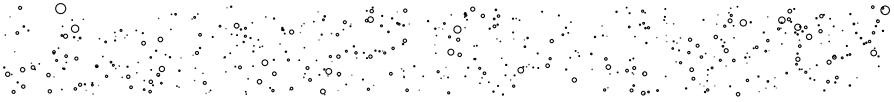
Die relevanten Fragen, die im Rahmen jeder Missionsplanung Gegenstand sein können, sind die folgenden:

- Bei gegebener Trajektorie und gegebener Katalogdichte, wieviele Bilder müssen im Mittel verarbeitet werden, bis eine Poseakquisition gelingt?
- Bei gegebener Trajektorie und gegebener Katalogdichte, wie hoch ist die Wahrscheinlichkeit, dass innerhalb eines vorgegebenen Akquisitionsfensters (in Bildern) eine Poseakquisition gelingt?

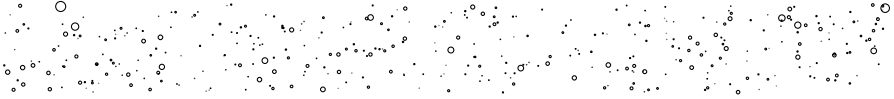
### 5.3.2 Testdaten und Testdefinition

Im Allgemeinen muss davon ausgegangen werden, dass der mitgeführte Kraterkatalog nicht dicht ist, da aufgrund der Eigenschaften des Kartierungsprozesses trotz an sich vorhandener Krater nur ein geringer Anteil von ihnen im Katalog verzeichnet sein kann. In der Konsequenz bedeutet das, dass die Akquisitionsperiode abhängig ist von dieser Dichte und mit sinkender Dichte länger werden wird. Ist eine optimale Abdeckung durch den Katalog gegeben, wird die Lösung (bei kleinen Messfehlern der Ellipsen) annähernd instantan sein. Sind kaum bekannte Krater zu sehen, sind möglicherweise viele Bilder zu verarbeiten und eventuell vergehen sogar mehrere Bildaustauschperioden, bevor eine Lösung vorliegt.

Die instantane Erfolgswahrscheinlichkeit für einen breiten Parameterbereich wurde sowohl durch vollsynthetische Testdaten im Kapitel 4 untersucht,



(a) Zufällige Kraterverteilung auf der Geländeebene



(b) Katalog zur Kraterverteilung mit Dichte 3/4



(c) Katalog zur Kraterverteilung mit Dichte 1/2



(d) Katalog zur Kraterverteilung mit Dichte 1/4

Abbildung 5.12: Eine Verteilung synthetischer Krater auf planem Testgelände für Tests der mittleren Akquisitionsperiode mit abgeleitetem Katalog in mehreren Dichtestufen

wo ein vollständiger Katalog vorlag, als auch im Abschnitt 5.2 dieses Kapitels, wo nur ein unvollständiger Katalog zur Verfügung stand (und noch weitere unquantifizierbare Fehler durch die Bildverarbeitung eingingen). Für die Bestimmung eines Basiswertes für die mittlere Akquisitionsperiode wird in diesem Abschnitt wieder ein vollsynthetischer Datensatz generiert, ohne Bildsimulation. So ist sichergestellt, dass einerseits das Störverhalten der Bildverarbeitung nicht in die Ergebnisse eingeht und andererseits, dass eine „unendliche“ Trajektorie zur Verfügung steht, um ausreichend Messungen durchführen zu können. Der Katalog wird in verschiedenen Dichtestufen in die Navigationsalgorithmen eingegeben, um die hierdurch bedingte Variabilität der Erkennungsperiode quantifizieren zu können. Abb. 5.12 zeigt eine solche zufällige Verteilung mit mehreren daraus abgeleiteten Katalogen abnehmender Dichtestufen.

### 5.3.2 Testdaten und Testdefinition

---

Insgesamt wird eine Anzahl von Tests durchgeführt werden, wobei jeweils von Beginn einer simulierten Trajektorie über ein synthetisches Kraterfeld, wie in Abb. 5.12a exemplarisch abgebildet, die Anzahl verarbeiteter Bilder gemessen wird bis zur erfolgreichen Poseakquisition. Der Referenztest, währenddessen die Akquisitionsperiode für vollständigen Bildaustausch gemessen wird (also die Position der Kamera auf der Testtrajektorie in ganzzahligen Vielfachen der Bildaustauschperiode inkrementiert wird), deckt den identisch unabhängig verteilten Messfall ab. Für sinkende Katalogdichte lässt er zudem Aufschluss über das Wachstum der mittleren Akquisitionsperiode zu. Für eine feste Katalogdichte, aber zunehmend höhere Messfrequenzen unterhalb einer Bildaustauschperiode, wird dann das Verhalten der mittleren Akquisitionsperiode unter teilweise korrelierten Messungen betrachtet.

Die Daten und Parameter der Tests werden dabei wie folgt variiert bzw. festgelegt:

**Verteilung der Testkrater in Ort und Radius** — Die Testkrater selbst werden wieder zufällig gezogen. Die Verteilung ihrer Radien entspricht der auf Seite 144 gewählten Exponentialverteilung und die Verteilung der Mittelpunkte der Gleichverteilung im Rechteck, das durch das Sichtfeld der Kamera auf dem simulierten Gelände überstrichen wird.

**Kraterkatalog** — Der Katalog wird so aus den generierten Testkratern erzeugt, dass er über die Trajektorie hinweg gleichförmig ausgedünnt wird mit fester Zieldichte  $\leq 100\%$ . Die Parameter der Verteilungen von Radien und Mittelpunkten ändern sich so nicht. Auf diese Weise wurde der in den Abbildungen 5.12b, 5.12c und 5.12d gezeigte Katalog erzeugt aus den darüber in Abb. 5.12a abgebildeten Testkratern.

**Kamera-Elevation und -Flughöhe** — Zusätzlich zur Trajektorienabfolge der synthetischen Ellipsen noch die Kameraposenparameter wie in Abschnitt 4.3 zu variieren, wäre rechnerisch nicht mehr tragbar. Anhand der in Abb. 5.11 auf Seite 203 und Tabelle 5.1 auf Seite 204 ermittelten Optimalwerte wird daher ein fester Winkel von  $45^\circ$  gewählt. Als Höhe der Kamera über dem simulierten Gelände wird wiederum der Wert gewählt, der sicherstellt, dass das simulierte Bild bis zu den Seitenrändern ausgefüllt ist.

**Simulierte Bahngeschwindigkeit** — Die simulierte Downrange-Geschwindigkeit der Kamera wird so gewählt, dass im Referenzlauf jeweils eine ganze Bildaustauschperiode vergeht zwischen zwei synthetischen Aufnahmen,

und dann zur Bewertung der Wirkung der Korrelation zwischen überlappenden Aufnahmen jede Bildaustauschperiode in eine Abfolge von mehreren Einzelbildern unterteilt wird.

**Testumfang** — Es werden pro Test eine feste Anzahl von 1000 Einzelmessungen (verarbeiteten simulierten Bildern) durchgeführt.

**Störung der synthetischen Ellipsen** — Der Test wird nur für ungestörte synthetische Ellipsen durchgeführt, da hier primär der Effekt der teilweisen Korrelation der Eingangsdaten und die Wirkung der variablen Katalogdichte von Interesse ist.

Jeder hierdurch beschriebene Einzelttest umfasst die Generierung einer Kraterverteilung und dazu eines Kataloges bestimmter festgelegter Dichte.

### 5.3.3 Auswertungsmethodik

Ziel dieser Untersuchungen ist, auf Basis der Testergebnisse überprüfbare Vorhersagen für die Performance während der im folgenden Abschnitt durchgeführten TRON-Labortests zu konstruieren. Dazu werden zunächst die in der Referenzmessreihe (jeweils volle Bildaustausche) entstandenen Daten bewertet und die zulässigen Schlüsse bezüglich der eingangs in diesem Abschnitt gestellten Fragen gezogen. Daraufhin wird geklärt werden, inwiefern der Übergang zu unvollständigen Bildaustauschen diese Ergebnisse beeinflussen, denn diese bilden die Bedingungen der TRON-Tests erheblich besser ab.

Es darf vermutet werden, dass sich für voneinander unabhängige Messungen die Verteilung der Akquisitionsergebnisse näherungsweise als POISSON-Prozess ausprägt und die Wartezeiten (in Vielfachen einer Bildaustauschperiode) zwischen den Ereignissen exponentialverteilt sind mit einem bestimmten Erwartungswert. Dieser Modellzusammenhang zwischen den unabhängigen Testreihenparametern und den Erwartungswerten (bzw. Mittelwerten) der resultierenden Wartezeitenverteilungen soll gefunden werden, um dann mit der gebotenen Vorsicht die gesuchten Aussagen über die mittlere Akquisitionsperiode abzuleiten und diese in den unter-Bildaustauschperiodenweise abgetasteten Vergleichsmessreihen möglichst zu bestätigen. Eine solche Bestätigung würde einen Hinweis darauf darstellen, dass aus den Referenztests hergeleitete Modellvorhersagen auch für den Fall der kleinen Bildinkremente Gültigkeit behalten und das Bildaustauschinkrement (in Einheit Bildaustauschperioden) zwischen zwei Messungen keinen signifikanten Einfluss hat.

Insbesondere wird auch untersucht werden, ob sich ein funktioneller Zusammenhang zwischen den in den Tests variierten Parametern Katalogdichte und Bildaustauschinkrement benennen lässt, um eine anhand der TRON-Tests überprüfbare Vorhersage der Wartezeit machen zu können.

### 5.3.4 Ergebnisse

Eine Staffelung der Katalogdichte  $d$  in Viertelschritten von  $1/4$  (dünner Katalog) bis  $1$  (voller Katalog) und Messung der Wartezeit  $w_a$  (in Bildern) zwischen aufeinanderfolgenden Akquisitionseignissen führt zu den Wartezeitenverteilungen, deren Histogramme in Abb. 5.13 tabelliert sind. Hierbei wurde jeweils die synthetische Trajektorienposition um genau den Downrange-Schritt inkrementiert, der einen vollständigen Bildaustausch erzeugte.

Die Einzelmessungen dieser Messreihe sind daher im oben besprochenen Sinne unabhängig voneinander und es ist daher zulässig, durch Bootstrap über das arithmetische Mittel der erzeugten Wartezeitenverteilung das jeweils gestrichelt eingezeichnete 95%-Konfidenzintervall  $E_{95\%}(w_a)$  des Erwartungswertes zu schätzen. Für jedes arithmetische Mittel  $\bar{w}_a$  der realisierten Wartezeitenverteilungen lässt sich auch jeweils eine Exponentialverteilung

$$f(w_a) = \lambda e^{-\lambda w_a} \quad (5.1)$$

mit  $\lambda := 1/\bar{w}_a$  angeben, wie ebenfalls in den Unterabbildungen als dicke schwarze Kurven eingezeichnet. Trägt man die Erwartungswerte und Mittelwerte über die Katalogdichte ab bzw. betrachtet die Verteilungsfunktionen dieser Exponentialmodelle, ergibt sich die Abb. 5.14. In Abb. 5.14a ist ebenfalls die Ausgleichsfunktion über die Stichprobenmittelwerte eingezeichnet:

$$\bar{w}_a^{\text{mdl}}(d) = 0.7402d^{-1.7766}, \quad (5.2)$$

wobei  $d$  wieder für die Dichte des Kataloges steht. Abb. 5.14b zeigt die kumulativen Verteilungsfunktionen der in den Einzelgraphen der Abb. 5.13 eingezeichneten Exponentialverteilungen zu den jeweiligen Wartezeiten-Mittelwerten.

Bezeichne  $\delta_i$  das Bildaustauschinkrement zwischen zwei Aufnahmen in der Einheit *Bildaustauschperiode* (bisher galt also  $\delta_i = 1$ ). Behält man einen Katalog voller Dichte bei, verringert aber sequentiell die Inkremente  $\delta_i$  zwischen zwei Aufnahmen, erhält man die Graphen der Abb. 5.15. Da bei Überlappung der Bilder wie oben besprochen die Messungen nicht mehr als identisch unabhängig verteilt betrachtet werden können, ist das oben berechnete Konfidenzintervall für diese Daten bedeutungslos. Der Stichprobenmittelwert  $\bar{w}_a$  ist



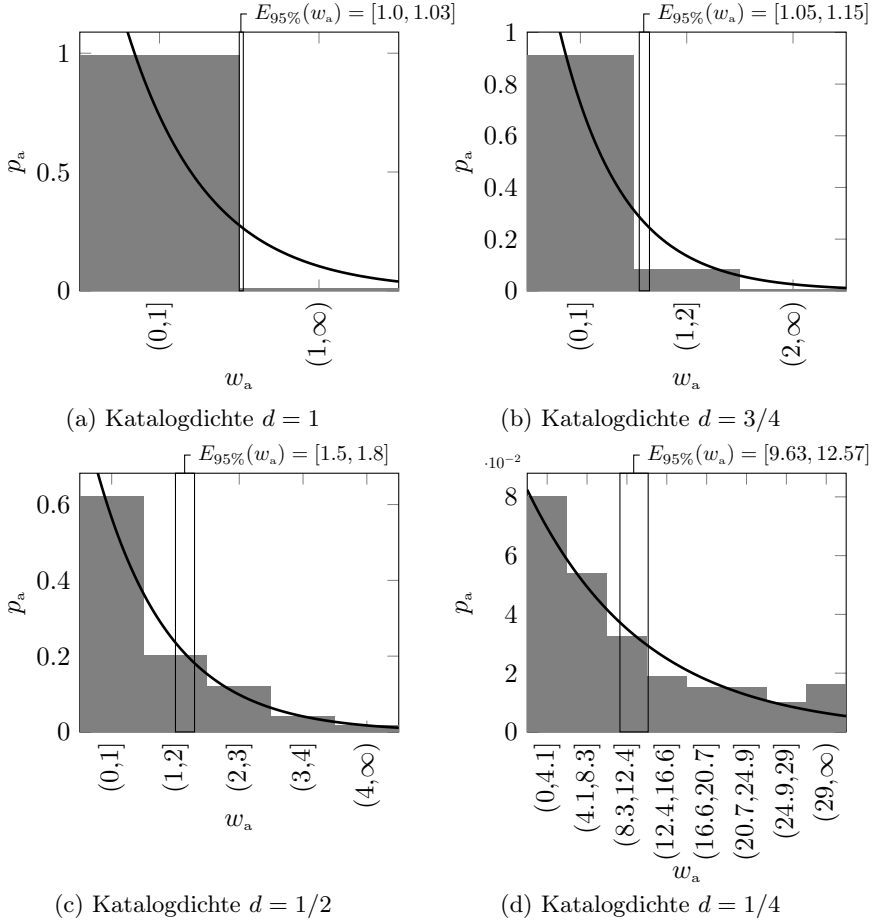
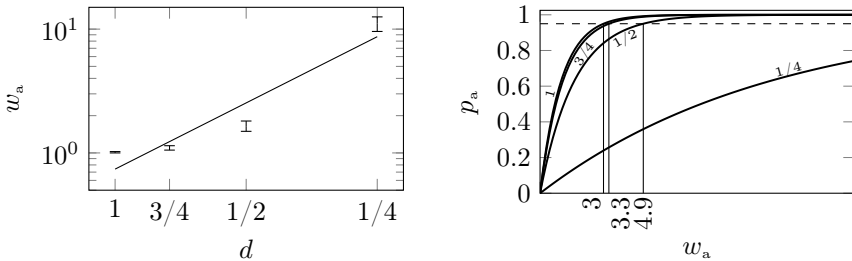


Abbildung 5.13: Wartezeitenverteilung in Bildern für vollständigen Bildaustausch und verschiedene Dichtestufen des Kataloges mit Exponentialverteilung zum arithmetischen Mittel und Konfidenzintervall zum Erwartungswert

### 5.3.5 Auswertung und Fazit



(a) Stichprobenmittelwerte-Ausgleichsfunktion und 95%-Konfidenzintervalle der Erwartungswerte der Wartezeiten in Abhängigkeit von der Katalogdichte (log-log) (b) Vollständige Verteilungsfunktionen der Wartezeiten zu den Stichprobenmittelwerten mit 95%-Quantilen

Abbildung 5.14: Verhalten der Erwartungswerte und 95%-Quantile der Wartezeitenverteilungen in Abhängigkeit von der Katalogdichte

aber markiert. Diese realisierten Mittelwerte sind in Abb. 5.16 in Abhängigkeit von den verwendeten Bildaustauschinkrementen dargestellt.

### 5.3.5 Auswertung und Fazit

**Referenzmessreihe** — Die der Referenzmessreihe zugrundeliegenden Annahmen bzw. Argumente waren, zusammengefasst:

- Es erfolgt ein vollständiger Bildaustausch zwischen jeweils aufeinanderfolgenden Messungen
- Über den gesamten Test sind die Parameter Katalogdichte, Gelände, Beleuchtung und Sichtbedingungen konstant
- Die Akquisitionsergebnisse prägen sich als homogener POISSON-Prozess aus

Der letzte Punkt impliziert, dass die Wartezeiten zwischen den jeweiligen Akquisitionsergebnissen exponentialverteilt sind. Durch Vergleich des Histogramms der Abb. 5.13a zur Exponentialkurve lässt sich, für einen dichten Katalog (jeder erkannte Krater ist im Katalog verzeichnet), diese Vermutung bestätigen. Auch mit abnehmender Katalogdichte bleibt die Approximationsgüte des Exponentialverteilungs-Modells größtenteils erhalten, wie in den

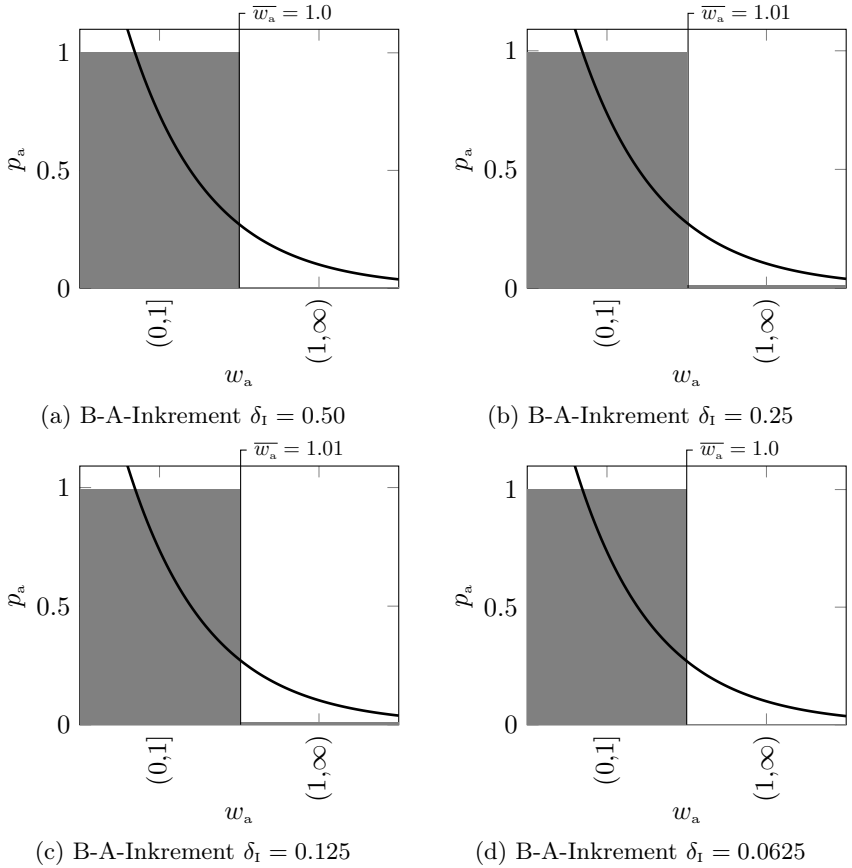


Abbildung 5.15: Wartezeitenverteilung in Bildern für variablen Bildaustausch und konstanter Dichte des Kataloges mit Exponentialverteilung zum arithmetischen Mittel

Abbn. 5.13b und 5.13c erkennbar, lediglich im Fall des dünnsten Kataloges in Abb. 5.13d scheinen unmodellerte Effekte eine Rolle zu spielen, die die tatsächliche Wartezeitenverteilung nach rechts verzerren. Es ist vor allem der zu dieser Verteilung gehörende Datenpunkt, der in Abb. 5.14a im Vergleich zur Interpolationsfunktion nach oben ausbricht.

In Abb. 5.14a zeigt sich, dass der Zusammenhang des Mittelwertes der War-

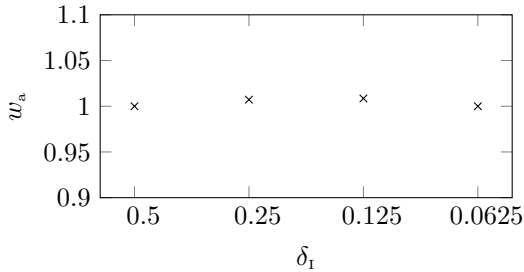


Abbildung 5.16: log-lin-Darstellung der Mittelwerte der Akquisitionswartezeiten in Abhängigkeit vom Bildaustausch zwischen jeweils aufeinanderfolgenden Messungen

tezeit in der log-log-Darstellung dennoch nahezu linear über die abnehmende Katalogdichte wächst, also auch hier ein gewisser exponentieller Zusammenhang besteht. Abb. 5.14b zeigt hierzu die erwartete Rechtsstreckung der Verteilungsfunktionen. Ebenfalls eingezeichnet sind dort die 95%-Quantile zu den Modellen.

Mit diesen Ergebnissen lässt sich die konkrete Frage nach dem Erwartungswert der Wartezeit auf eine erfolgreiche Akquisition, also nach der *mittleren Akquisitionsperiode*, wie folgt beantworten: Sie entspricht dem Erwartungswert der Wartezeitenverteilung<sup>4</sup>. Beispielsweise lässt sich für den vollständigen Katalog (gemessen als Anteil der erkannten Krater, abzüglich der Fehlerkennungen, die im Katalog verzeichnet sind) den erhobenen Daten zufolge also eine erfolgreiche Akquisition innerhalb von einem bis 1.03 Bildaustauschen erwarten (vgl. Abb. 5.13a). Eine 95%tige Akquisitionswahrscheinlichkeit liegt im Modell bei ca. drei Bildern (vgl. Abb. 5.14b).

**Korrelierte Messreihe** — Die in der Abb. 5.15 tabellierten Ergebnisse der Tests mit Bildaustauschinkrement  $\delta_I < 1$  zeigen keine Rechtsstreckung der Exponentialkurve zum Mittelwert im Vergleich zur Abb. 5.13a. Abweichungen des Stichprobenmittelwertes der Wartezeitenverteilung in den Abbildungen 5.15b und 5.15c sind so klein, dass sie zufälligen Effekten zugerechnet werden können und nicht in einer Systematik zu suchen sind. Hierfür spricht auch, dass der Mittelwert der in Abb. 5.15d gezeigten Verteilung wieder zur Eins zurückkehrt.

---

<sup>4</sup>Diese Tatsache ist bekannt als das sog. „Wartezeitparadoxon“.

$\delta_I$	$\overline{w}_a$	Ref.-Test
1.0000	1.0102	✓
0.5000	1.0000	✗
0.2500	1.0072	✗
0.1250	1.0085	✗
0.0625	1.0000	✗

Tabelle 5.2: Mittelwerte der realisierten Wartezeiten für schrumpfende Bildaustauschinkremente mit Vergleich zur Modellvorhersage

Setzt man die Stichprobenmittel mit dem des Referenztests für diese Katalogdichte und den Vorhersagen des Modells (5.2) in Relation, erhält man die Tabelle 5.2. Man sieht hier, dass die Mittelwerte der Tests mit  $\delta_I < 1$  aus Abb. 5.16 beinahe identisch mit dem im Referenztest realisierten Vergleichswert sind.

Aus den Positionen der Stichprobenmittel in Abb. 5.15 und im Detail in Abb. 5.16 ist ersichtlich, dass das Bildaustauschinkrement  $\delta_I$  kaum Einfluss auf die mittlere Akquisitionsperiode hat. Als simples Vorhersagemodell für den Mittelwert der Wartezeiten ist also die Gleichung (5.2) daher genausogut geeignet wie im Fall  $\delta_I = 1$ . Die Approximationstreue dieser Funktion ist nicht sonderlich hoch, wie man in Abb. 5.14a erkennt. Mit Hinblick auf die nicht berücksichtigten und letztenendes ohnehin unquantifizierbaren durch die Bildverarbeitung eingebrachten Fehler, lohnt es aber nicht, hier ein komplexeres Modell höherer Ordnung zu suchen. Dies wird auch während der Auswertung der TRON-Tests in Abschnitt 5.4 deutlich werden.

Da in diesen Testreihen durchweg nur jeweils einige wenige der Akquisitionsversuche fehlschlagen, kann man nur schließen, dass eine reine Verkürzung des Bildaustauschinkrements die Akquisitionsperiode zumindest nicht verlängert. Die kombinierte Auswertung von Testreihen mit nicht vollständigem Katalog und sehr kleinen Bildaustauschinkrementen  $\delta_I$  ist Gegenstand des folgenden Abschnitts.

## 5.4 TRON-Labortests

Die Tests des vorangehenden Abschnittes basieren vollständig auf synthetischen Daten. Die in Abschnitt 5.3.4 gesammelten Ergebnisse bilden wiederum nur die wohldefinierten Störungen der synthetischen Ellipsen ab. Wie man an den Auswertungen der systematischen Abdeckungstests des Kamera- und

#### 5.4.1 Versuchsaufbau und Trajektoriendefinition

---

Beleuchtungs-Posenraumes im Abschnitt 5.2 aber ablesen konnte, ist nicht das gesamte Störungsverhalten durch diese Art von Eingangsfehlern erklärbar.

Das konkrete exponentielle Vorhersagemodell (5.2) auf Seite 210 basiert ebenfalls rein auf der (bildweisen, effektiven) Katalogdichte  $d$ , gemessen als der Anteil der im Bild erkannten Krater, die auch im Katalog verzeichnet sind.

Zur Bewertung dieses Modells wird daher die Testreihe zur Akquisitionspanne auf realistische Eingangsdaten erweitert im Rahmen eines Laborversuchs im TRON-Labor des DLR-Instituts für Raumfahrtssysteme in Bremen. Es steht eine Fräsung des Mondoberflächenmodells aus Abb. 5.1 zur Verfügung, das im Labor ca. 400km mal 200km Mondoberfläche in einem Maßstab von 1:125.000 abbildet (vgl. Abb. 5.17). Für Details des zum Zeitpunkt der Fertigstellung dieser Arbeit größtenteils noch aktuellen Ausbauzustandes des Labors, siehe (Krüger u. a., 2014).

#### 5.4.1 Versuchsaufbau und Trajektoriendefinition

Der Technische Versuchsaufbau steht wie folgt:

- Kameraführung: Sechs räumliche Freiheitsgrade durch einen KUKA KR16-Roboter
- Geländemodell: 3.92m×1.96m mit Geländedynamik von ca. 20cm, durch Fräsung gefertigt aus Kaguya Stereo-DEMs (vgl. Abb. 5.17)
- Kamera: Allied Vision GC1380H mit Optik Schneider-Kreuznach Cinegon 1.9/10, Abgebildeter horizontaler und vertikaler Öffnungswinkel ca. 36°
- Innere Kamerakalibration: Durch Software *CalLab* (Strobl u. a., nodate)
- Bildentzerrung: Durch OpenCV Version 2.3

Die Trajektorien der Kamera sind einem Teil eines Descent-Orbits nachmodelliert, beginnend in ca. 123km Flughöhe über dem Mondreferenzradius von 1.737.400m und endend in ca 113km Höhe. Da nur ein physisches Geländemodell zur Verfügung steht, wird mit ihm das unter der Kamera sichtbare Gelände „gekachelt“, d.h. das Modell wird effektiv mehrfach überflogen durch die Kamera. Die modellierte Trajektorie ist äquatorial und beginnt tagwärts am Terminator. Das erste Bild weist also eine minimale Beleuchtungs-Elevation von ca. 5° auf, das letzte eine maximale von ca. 50°. Es wird damit der in den Tests zur instantanen Erfolgsrate ermittelte Arbeitsbereich des Algorithmus gut abgedeckt.



Abbildung 5.17: Versuchsaufbau für Trajektorientests im Labor TRON am Institut für Raumfahrtssysteme des DLR in Bremen

*Bildquelle* Deutsches Zentrum für Luft- und Raumfahrt

Weiterhin liegen optische Achse und Lampenachse immer in einer Ebene, und zwar so, dass die Beleuchtung direkt „von vorn“ kommt. Dies ist ein potentiell pathologischer Fall der bildverarbeitenden Verfahren, da auf diese Weise oft der ausgeleuchtete Teil der Krater kaum noch eingesehen werden kann und die Erkennungsheuristiken oft nicht mehr anwendbar sind. Außerdem entstehen Linsenartefakte, die als störende Strukturen der Algorithmen beeinflussen. Diese maximal unvorteilhafte Testkonfiguration wurde gewählt, damit die erreichten Ergebnisse plausibel als Worst Case im Sinne einer unteren Schranke verstanden werden können.

Die Konfiguration der Kamerakoordinatenachsen mit den ersten und letzten Downrange-Sichtfeldgrenzen für einen solchen Überflug ist in Abb. 5.18

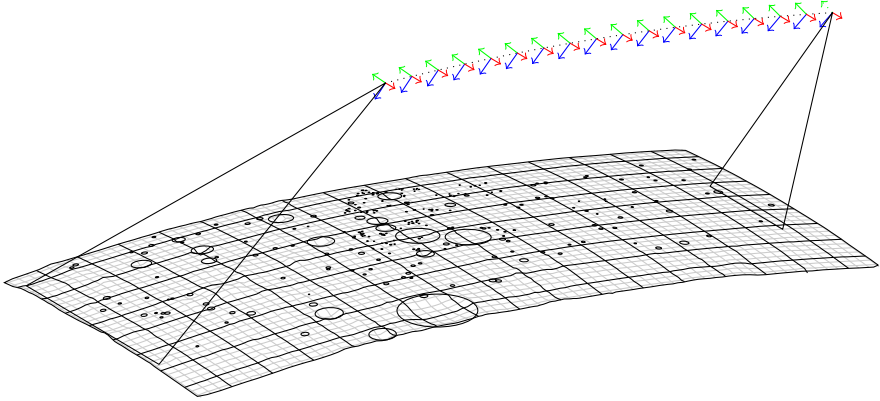


Abbildung 5.18: Trajektorie mit Downrange-Sichtfeldgrenzen für Tests zur Bestimmung der mittleren Akquisitionsperiode im TRON-Versuch, Kamera-Achsen in Reihenfolge rot, grün und blau, Sichtfeld  $36^\circ$ , nur einige Kameraposen abgebildet

dargestellt. Diese Grenzen sind so gewählt, dass die Bildaufnahme mit einem gefüllten Bild beginnt und ended, sobald der entfernte Rand des Gelände-modells in das Sichtfeld kommt. Dadurch wird verhindert, dass unnatürliche Konzentrationen von Krater-Messpunkten im Bild auftreten. Der ebenfalls in der Abbildung dargestellte Kraterkatalog für das Modell wurde mittels einer für diesen Zweck geschriebenen Software manuell erstellt durch Markieren von Kratern im Quell-DEM, er enthält ca. 250 Einträge. Für die Tests wird ein wesentlich dichter Katalog verwendet, der mithilfe eines SLAM (Simultaneous Localization And Mapping)-Verfahrens aus Messungen der in dieser Arbeit verwendeten Kratererkennungsalgorithmen automatisch generiert wurde (Kudari, 2015).

Die Kamera rotiert während jedes Überfluges so um ihre  $x$ -Achse, dass der Kamera-Elevationswinkel im Mittel bei  $45^\circ$  bleibt. Dies liegt wieder im optimalen Bereich, der im Abschnitt 5.2 ermittelt wurde. Das Bildaustauschinkrement dieser Messreihe ist leicht variabel über die Trajektorie hinweg und liegt bei  $\delta_1 \approx 7.34 \cdot 10^{-3}$ .



## 5.4.2 Bilddaten

Es wurde auf diese Weise insgesamt 1442 Bilder aufgenommen, unterteilt in sieben Überflüge zu je 206 Bildern mit sequentiell ansteigender Beleuchtungselevation. Ein Auszug aus diesen Daten ist in Abb. 5.19 gezeigt.

Die Bilder wurden vor der Verarbeitung durch die in OpenCV 2.3 verfügbaren Routinen entzerrt, ausgehend von inneren Kameraparametern, die vorher durch die Software *CalLab* ermittelt wurden. Damit ist sichergestellt, dass kreisförmige Kraterränder auf Ellipsen im Bild projizieren und damit die Annahmen der Detektionsalgorithmen erfüllt sind.

Auf diese Weise sind im Gegensatz zu den synthetischen Tests des vorangegangenen Abschnittes die äußeren Bedingungen nicht vollkommen konstant. Im Sinne der Überprüfung der Belastbarkeit des Modells (5.2) der mittleren Akquisitionsperiode und aufgrund der ohnehin unquantifizierbaren und hochvariablen Fehler der Bildverarbeitung wird diese Einschränkung aber im Interesse höheren Realismus bewusst in Kauf genommen.

## 5.4.3 Auswertungsmethodik

Ausgehend vom Vorgehen in Abschnitt 5.3.3 wird wiederum die Verteilung der tatsächlich realisierten Wartezeiten zwischen aufeinanderfolgenden Akquisitionen bewertet. Rückschlüsse aus den Wartezeitverteilungen auf den zugrundeliegenden Ereignisprozess bleiben aber leider unzulässig, da die Messungen erneut wegen Überlappung der Bilder korrelieren und für eine Auswertung, die mit vollen Bildaustauschen die Referenzmessreihe aus Abschnitt 5.3 abbildet, schlicht „zu wenig Gelände“ zur Verfügung steht.

Weiterhin wird darüber hinaus wegen der in den TRON-Tests nicht konstanten Kraterkatalogdichte eine detailliertere (nämlich akquisitionseignisweise) Betrachtung der Abweichung vom Wartezeitmodell notwendig. Um eine Vorhersage für die Länge einer Wartezeit mit dem tatsächlichen Akquisitionseignis überhaupt sinnvoll in Relation setzen zu können, muss dabei angenommen werden, dass im Warteintervall vor jedem solchen Ereignis die Katalogdichte konstant war. Hierzu wird das arithmetische Mittel der während der erfolglosen Versuche beobachteten Katalogdichten als Eingang für das Wartezeitmodell am Akquisitionseignis gewählt, mit dem sein Wert verglichen werden soll.

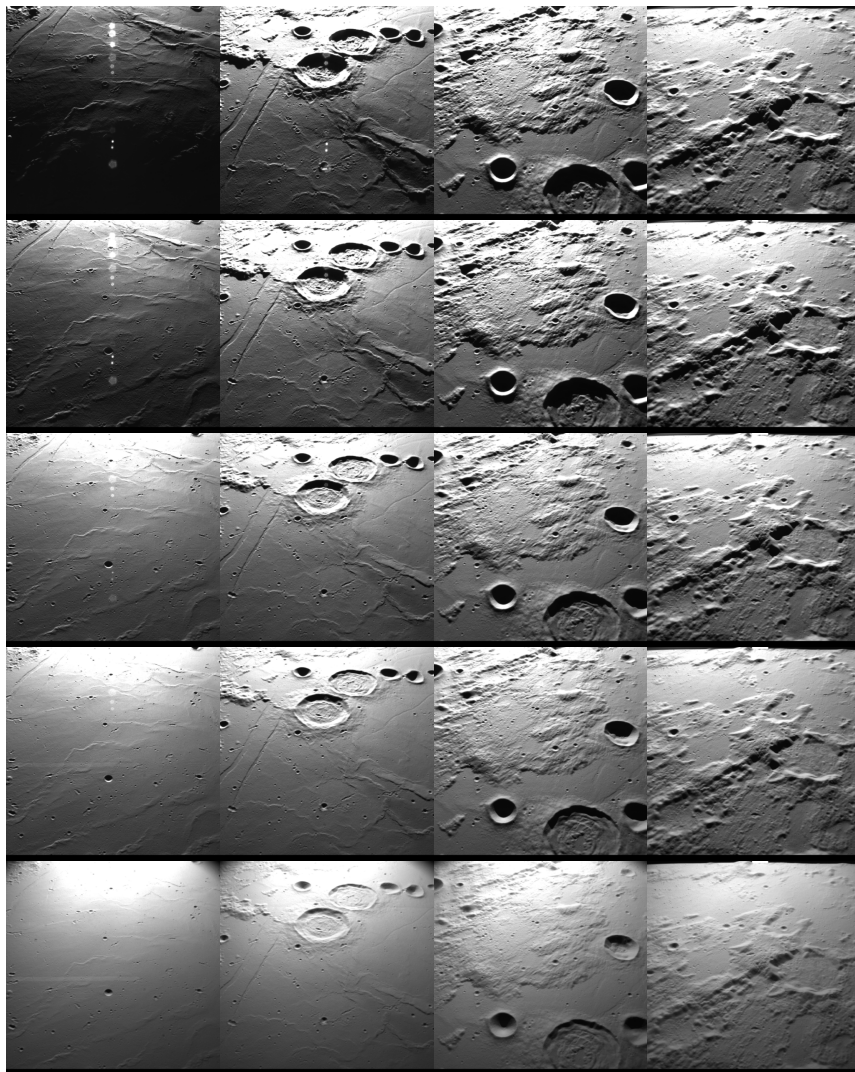


Abbildung 5.19: Auszug aus dem aufgenommenen Bildmaterial für die TRON-Tests mit Trajektorienverlauf von oben nach unten und ansteigender Beleuchtungs-Elevation von links nach rechts

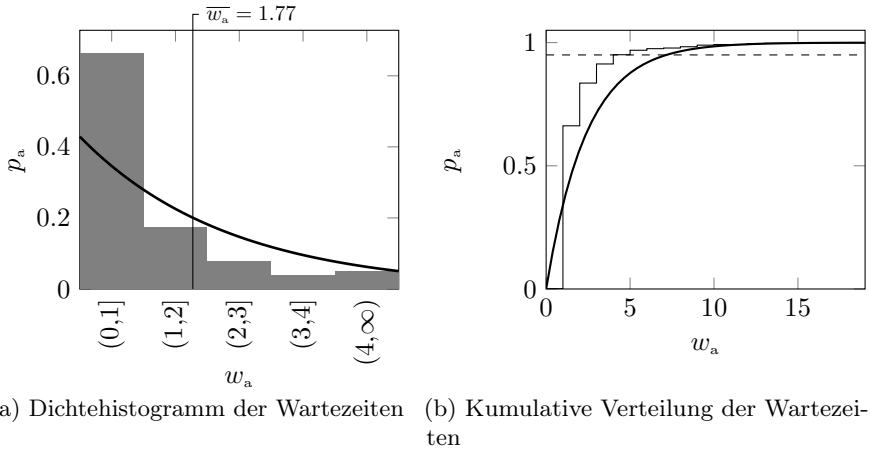


Abbildung 5.20: Wartezeitenverteilungen im TRON-Test und Exponentialmodell zum Stichprobenmittelwert.

## 5.4.4 Ergebnisse, Auswertung und Fazit

Die den Verteilungshistogrammen aus Abschnitt 5.3.4 entsprechende Wartezeitenverteilung ist in Abb. 5.20a gezeigt. Über den eingezeichneten Stichprobenmittelwert der Wartezeiten ist erneut ein Exponentialmodell gefittet, dessen kumulative Verteilungsfunktion im Vergleich mit der empirischen kumulativen Verteilung in Abb. 5.20b gezeigt ist.

Die empirische Verteilung stimmt hier sichtbar nicht mit dem Modell überein. Der Grund hierfür liegt in Ausreißern, die dazu führen, dass das Modell deutlich rechtsgestreckt wird. Insbesondere im letzten, rechtsoffenen, Akkumulator des Histogramms kann man einen erheblichen Schwanz der Verteilung beobachten. Es kann also über die Bedenken aufgrund der Korrelation der Einzelmessungen hinaus nicht mehr davon ausgegangen werden, dass die tatsächliche zu erwartende mittlere Akquisitionsperiode noch in der Nähe des Mittelwertes dieser Verteilung liegt. Der Ursprung der Ausreißer liegt hauptsächlich darin, dass der Akquisitionereignisprozess nicht mehr homogen ist, da die effektive Katalogdichte über die Trajektorie stark variiert. Dargestellt ist genau das in der oberen Achse der Abb. 5.21.

Die Ursache der dort sichtbaren periodischen Struktur in der Katalogdichte ist in den variablen Beleuchtungsbedingungen zu suchen: Jeweils zu Beginn

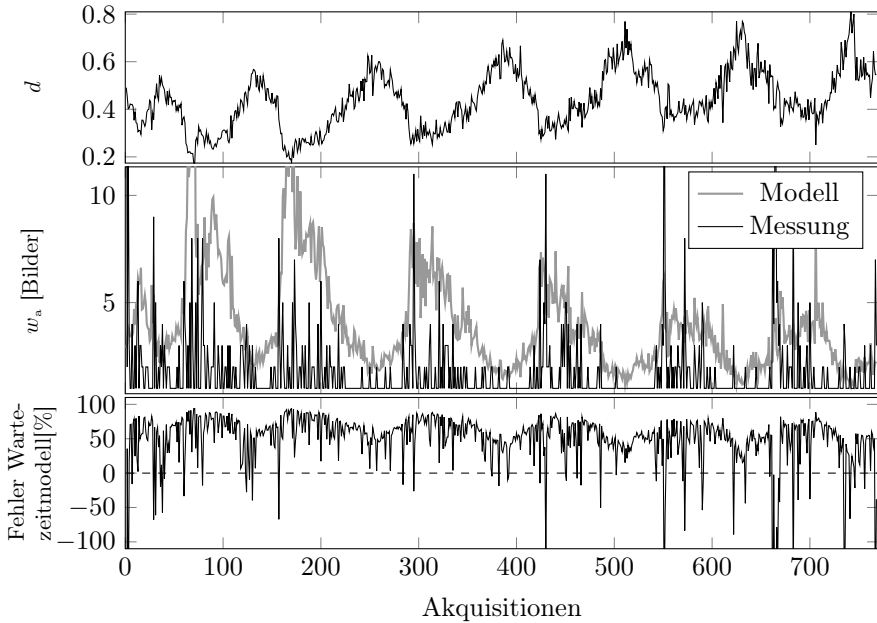


Abbildung 5.21: TRON-Test: Gemessene Wartezeiten vor den Akquisitionsergebnissen im Vergleich zu ihren Vorhersagen über die variable Katalogdichte mit Betrachtung der relativen Modellfehler.

eines der sieben Überflüge des Geländes ist die Beleuchtungselevation minimal, jeweils zum Ende hin maximal. Die Software, die den Katalog erstellte, tat dies aber (realistischerweise) für eine global mittelhohe Referenzbeleuchtungselevation und verzeichnete damit andere Krater als die, die jetzt im Test erkannt werden. Auch ist eine Gesamttendenz der effektiven Katalogdichte nach oben zu beobachten. Diese ist mit derselben Begründung in der sukzessive über die Einzelüberflüge steigenden Beleuchtungselevation zu suchen (das Szenario ist ein Flug „in den Morgen hinein“).

Die mittlere Achse der Abb. 5.21 zeigt die Gegenüberstellung der tatsächlich realisierten Akquisitionswartezeiten mit der Vorhersage durch das Modell (5.2), mit dem berechneten relativen Fehler der Vorhersage in der unteren Achse. Auffällig ist, dass die Modellvorhersage für mittlere und höhe-

re Katalogdichten die tatsächliche Akquisitionsperiode relativ konstant um 50% bis 90% überschätzt, für niedrigste Katalogdichten aber eine teilweise sehr akkurate Schätzung gibt. Auch erscheint die Modellkurve in gewissem Sinne „stetiger“ als die Kurve der realisierten Akquisitionsperioden, die ab einer Katalogdichte  $d \geq 0.4$  ca. ein Sättigungsverhalten gegen die optimale Untergrenze von unterbrechungsloser Akquisition zeigt. Erklärbar ist dieses an sich sehr willkommene Verhalten nicht, es war nicht in den synthetischen Tests reproduzierbar. Daher lässt sich vermuten, dass es dort noch unmodellerte Aspekte gibt, die das Akquisitionsverhalten des Verfahrens in dieser vorteilhaften Weise gestalten.

Als Fazit lässt sich festhalten, dass das Modell für *a-posteriori* bekannte Katalogdichten eine konservative Schätzung liefert, die um ca. 5% der Akquisitionswarteintervalle unterschätzt.

Letztendlich sind aber die beobachteten Modellabweichungen nicht ausschlaggebend, denn um mittels des Modells für einen zufälligen Einschaltpunkt während einer Trajektorie eine *a-priori*-Vorhersage der zu erwartenden Akquisitionsperiode zu treffen, muss die effektive Katalogdichte geschätzt werden. In realistischen Missionsplanungs-Szenarios wäre sie nicht im oben bekannten (*a-posteriori*) Detailgrad bekannt und wäre auch kaum mit größerer Genauigkeit als ca. 20% abschätzbar, denn solange nicht alle physisch vorhandenen Krater im Katalog verzeichnet sind, lässt sich die Überschneidung von Katalog und Erkennungen nur approximieren.

Ein realistischerer Test der Vorhersagekraft des Modells wäre also die Performance des Verfahrens im Vergleich zum Modellwert zu einer konservativen und konstanten Schätzung der Katalogdichte. Hierfür sei das arithmetische Mittel  $\bar{d} = 0.4111$  der beobachteten Dichten  $d$  gewählt, was zu einer konstanten Vorhersage der Akquisitionsperiode von  $\bar{w}_a^{\text{mdl}}(\bar{d}) = 3.5911$  Bildern führt. Dieser Wert ist die Vorhersage für die Wartezeit, die für einen Einstieg an einer zufällig gewählten Stelle der Trajektorie im Mittel eingeplant werden muss, bevor eine Akquisition erreicht wird.

Zunächst ist der statische Vergleich der tatsächlichen Wartezeiten mit der Vorhersage zu dieser mittleren Katalogdichte in Abb. 5.22 gezeigt. Hier ergibt sich, wie zu erwarten, eine höhere Anzahl an grob unterschätzten Akquisitionsperioden, die, wie aus Abb. 5.21 bekannt, wiederum am Ende von Intervallen extrem niedriger effektiver Katalogdichte liegen.

Wie hoch ist jedoch die mittlere Akquisitionsperiode? Aufgrund der in Abb. 5.20 sichtbaren Abweichung der globalen Wartezeitverteilung vom Exponentialmodell lässt sich der Mittelwert der wahren Wartezeiten nicht in der Nähe des Mittelwertes der Akquisitionswartezeitintervalle vermuten, denn es

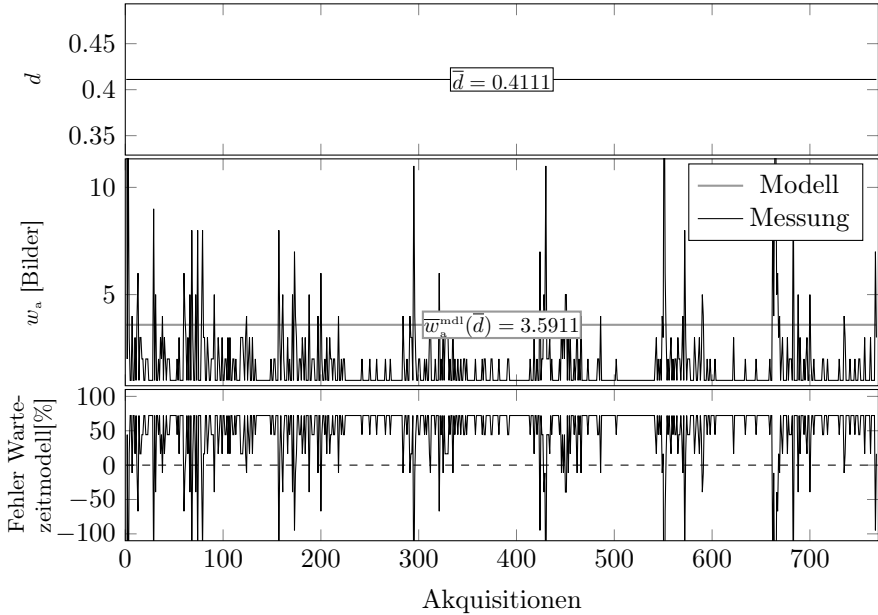


Abbildung 5.22: TRON-Test: Gemessene Wartezeiten vor den Akquisitionsergebnissen im Vergleich zu ihren Vorhersagen über die konstante Katalogdichte mit Betrachtung der relativen Modellfehler.

gibt erheblich mehr sehr lange Wartezeitintervalle, als von der Exponentialverteilung modellierbar. Damit steigt natürlich auch die Chance, bei zufälligem Einschaltzeitpunkt sich genau in einem solchen Intervall wiederzufinden.

Stehe  $t_a = (t_{a1}, \dots, t_{aN_a})$  für die in Bildern gemessenen Akquisitionszeitpunkte. Modelliert man die zufälligen Einschaltzeitpunkte  $t$  als diskrete Gleichverteilung auf den ganzen Zahlen im Intervall zwischen dem Einstiegs- punkt vor dem ersten Bild der Testserie und dem Bild zur letzten erfolgreichen Akquisition,

$$t \sim \mathcal{U}\{0, 1, \dots, t_{aN_a}\}, \quad (5.3)$$

wobei jeweils mindestens ein Bild verarbeitet werden muss, ist die jeweilige

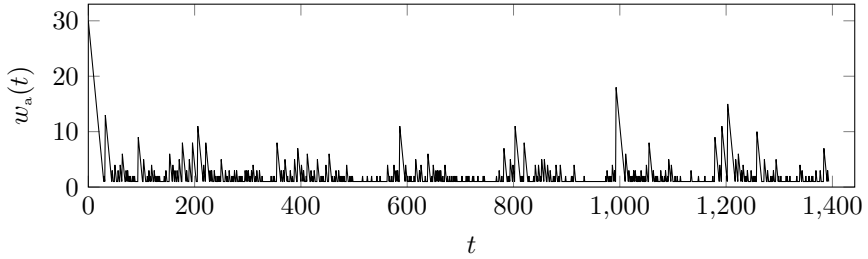


Abbildung 5.23: Bildweise Wartezeiten bis zum nächsten Akquisitionseignis im TRON-Test

Wartezeit gegeben durch

$$w_a(t) = t_{ai} - t \text{ für } i = \arg \min_j \{t_{aj} > t\}. \quad (5.4)$$

Der Wartezeitengraph über alle solche möglichen  $t$  ist in Abb. 5.23 gezeigt. Da alle  $t$  gleich wahrscheinlich realisiert werden, ergibt sich die mittlere Akquisitionsperiode durch das arithmetische Mittel

$$\frac{1}{t_{aN_a} + 1} \sum_{i=0}^{t_{aN_a}} w_a(i) = 2.4857 \text{ Bilder}. \quad (5.5)$$

Dieser Wert liegt unterhalb der Modellvorhersage von 3.5911 Bildern, die Strategie der Nutzung der mittleren Katalogdichte zur Vorhersage der mittleren Akquisitionsperiode beim inhomogenen Ereignisprozess ist also dieser Testreihe zufolge nicht unzulässig. Klar verwerfen muss man demgegenüber jetzt schon die Exponentialhypothese über die gemessenen Akquisitionswartezeiten aus Abb. 5.20, der dort ermittelte Wert lag bei 1.77 Bildern. Um eine robuste Aussage über Bestätigung oder Ablehnung dieses Ansatzes machen zu können, sollte man allerdings weitere Messreihen durchführen, sofern entsprechendes Datenmaterial zur Verfügung steht. Im Rahmen dieser Dissertation ist hier leider aufgrund des Mangels an solchem Material die Grenze der Verifizierbarkeit des Modells erreicht.

Eingangs wurde noch die Frage nach der Akquisitionswahrscheinlichkeit bei vorgegebenem Zeitfenster  $w$  und vorgegebener Katalogdichte  $d$  gestellt. Grundsätzlich lässt sich das Modell (5.2) zur Beantwortung dieser Frage verallgemeinern, denn leitet man für die Exponentialverteilungsdichte zum

#### 5.4.4 Ergebnisse, Auswertung und Fazit

---

<b><math>d</math> fest</b>	<b><math>w</math> fest</b>	<b><math>p_a^{\text{mdl}}</math></b>	<b><math>p_a</math> a-posteriori</b>
0.4111	1 Bild	0.2431	0.5509
0.4111	2 Bilder	0.4270	0.7382
0.4111	5 Bilder	0.7515	0.9110
0.4111	10 Bilder	0.9383	0.9713

---

Tabelle 5.3: Modellgeschätzte Akquisitionswahrscheinlichkeiten zu fester angenommener Katalogdichte und festem angenommenem Akquisitionszeitfenster

durch das Modell geschätzten Mittelwert die Verteilungsfunktion her, so erhält man eine geschätzte Akquisitionswahrscheinlichkeit  $p_a^{\text{mdl}}$  mit

$$p_a^{\text{mdl}}(d, w) = 1 - \exp\left(-\frac{w}{\bar{w}_a^{\text{mdl}}(d)}\right). \quad (5.6)$$

Da bei unbekanntem  $\bar{w}_a$  allerdings mit der approximationsfehlerbehafteten Modellvorhersage im Nenner gearbeitet wird, ist die Modellgüte von (5.6) tendenziell noch kleiner als die von (5.2). Dieses Modell würde daher nur eine sehr grobe Schätzung erlauben. Einige Beispielberechnungen sind in Tabelle 5.3 aufgelistet, zusammen mit den anhand der TRON-Messreihe und dem oben benutzten gleichverteilten Abtastungsansatz ermittelten empirischen Werte.

Man erkennt, dass die Akquisitionswahrscheinlichkeitsschätzungen durch das Modell (5.6) zu niedrig ausfallen in dem Maße, das der Überschätzung der mittleren Akquisitionsperiode durch das Modell (5.2) entspricht. Zur einer konservativen Ersteinschätzung erscheint das Modell daher nicht ungeeignet.

Als Fazit der oben gemachten Tests und Auswertungen lässt sich folgende Zusammenfassung geben:

- Die in Abschnitt 5.3 vorgeschlagene Hypothese eines exponentiellen Wartezeitmodells ist auf realistische Szenarien und die dabei realisierten Akquisitionsergebnisprozesse nicht anwendbar, denn es gibt eine unmodelliert große Anzahl sehr langer Akquisitionswartezeiten.
- Der Grund für diese Abweichungen ist primär in der inhomogenen effektiven Katalogdichte zu suchen, sekundär in den Abhängigkeitseffekten zwischen Messungen auf aufeinanderfolgenden Bildern.
- Es existieren weitere Effekte, insbesondere das unmodellierte Sättigungsverhalten aus Abb. 5.21. Aufgrund der großen Unsicherheit in der *a-priori*-Schätzung der unabhängigen Hauptgröße, der Katalogdichte, verspricht eine Verbesserung des Modells aber keinen Erkenntnisgewinn.



- Das Modell (5.2) liefert ausgehend von einer konservativen aber konstanten Schätzung der Katalogdichte eine gute Vorhersage des über die gesamte Trajektorie ermittelten mittleren Akquisitionsintervalls ausgehend von zufälligen Einschaltzeitpunkten.
- Eine sehr konservative Abschätzung der Akquisitionswahrscheinlichkeit zu einer gegebenen Katalogdichte und einem festen Akquisitionszeitfenster ist durch das Modell (5.6) möglich.



# Kapitel 6

## Fazit und Ausblick

### 6.1 Fazit

In der vorliegenden Arbeit ist detailliert ein Verfahren hergeleitet worden, das die eingangs auf Seite 7 formulierte zentrale Zielstellung erfüllen kann:

„Aus einem statischen Katalog von Mondkratern und einem einzigen oder wenigen nicht in besonderem zeitlichen Zusammenhang zueinander stehenden geeigneten Bildern der Mondoberfläche, identifiziere ohne weitere Information ausreichend viele der abgebildeten Krater, um auf die Position und Lage der aufnehmenden Kamera rückschließen zu können. Die Anforderungen an die Lage der Kamera zur Oberfläche und an die externe Beleuchtung der Oberfläche sollen dabei minimal sein.“

Es kombiniert die in Kapitel 2 beschriebene und erweiterte Methode zur Detektion von Kratern in einem Bild der Mondoberfläche mit den danach im Kapitel 3 entwickelten Algorithmen, um daraus die gewünschte Identifikation der gesehenen Krater zu erreichen. Die dazu gefundenen Lösungen übertreffen den dem Autor bekannten Stand der Forschung auf verschiedenen Gebieten.

Der bildverarbeitende Teil des Algorithmus beinhaltet dabei konkret folgende Fortschritte gegenüber dem State of the Art:

- Die vorgestellte Methode zur Suche nach stabilen Bildkontrastgebieten ist bezüglich der intern zu verfolgenden Metadaten erheblich schlanker

als sein nächster Verwandter, das MSER-Verfahren, und um Größenordnungen simpler als das von Spigai, Clerc und Vincent Simard Bilodeau verwendete Verfahren. Darüber hinaus muss das Bild selbst nur einmal sequentiell gelesen werden und es sind keine Finite-Differenzen-Karten notwendig für Gradientenberechnungen, wie in den kantendetektierenden Verfahren der Art von Cheng, Johnson u. a.

- Die Lichtrichtung im Bild wird mittels des vorgestellten Verfahrens über den Nachbarschaftsgraphen der stabilen Kontrastgebiete ohne äußere Hilfsmittel abgeleitet. Dem Autor ist kein anderes Verfahren bekannt, das diese Richtungsbestimmung in solcher Einfachheit beherrscht, insbesondere ohne erneuten Zugriff auf die Bilddaten.

Die Methode der Freiform-Modellbildung über die rekonstruierten räumlichen Kraterrendnormalen ermöglicht das Übertreffen bisheriger Krateridentifikationsverfahren in den folgenden Punkten:

- Die Fähigkeit zur Modellierung nicht-planarer Oberflächen erlaubt die Identifikation von Kraterkonstellationen, die durch das Verfahren von Ansar und Cheng aufgrund der Beschränkung auf koplanare affine Invarianten nicht mehr abgedeckt werden.
- Die Modellierung von Oberflächen mit anisotropen Krümmungsfeldern übertrifft das Vermögen der von D. S. Wokes und S. M. Wokes vorgestellten Methode zur Nadirbestimmung, die nur strikt kugelförmige Gelände verarbeiten kann.
- Beherrschung der perspektivischen Darstellung von nicht-planaren Oberflächen erweitert die Fähigkeit der Identifikation über die des Verfahrens von Hanak hinaus, das auf eine strikt nadirblickende Kamerakonfiguration und vernachlässigbare Tiefendynamik angewiesen ist.

Es wurde eine Referenzimplementierung für die vorgestellten Methoden erstellt. Die Modellbildungs- und Identifikationsalgorithmen wurden im Kapitel 4 sorgfältig isoliert untersucht auf die Einzelwirkungen von modellierten punktuellen Fehlern der Kraterdetektion, wie sie durch Bildstörungen auftreten, und die Wirkung von modellierten systematischen Fehlern in deren Ergebnissen, wie sie durch Fehlkalibration von Kameras oder mangelhaftes Alignment der Sensoren am Raumfahrzeug erzeugt werden können. Es zeigte sich, dass selbst große Eingangsstörungen dieser Art meist keine katastrophale Störung der Ausgänge erzeugen und regelmäßig trotzdem eine Krateridentifikation möglich ist.

Tests im geschlossenen Kreis mit der Kraterdetektion in Kapitel 5 belegen die Robustheit über einen großen Bereich von Parametern der Beleuchtung eines Geländes und der Lage der aufnehmenden Kamera zum Gelände. Bezüglich der Robustheit gegenüber extremen Beleuchtungsbedingungen werden dabei in der Literatur anzufindende Toleranzbereiche erreicht oder übertroffen. Die Robustheit gegenüber sehr niedrigen Kamera-Elevationen ist schon methodenbedingt größer als die der Konkurrenzverfahren von Hanak oder Cheng, Johnson u. a. in dieser Nische.

Da das hier entwickelte Verfahren der Erst- oder Neuinitialisierung eines Zustandsschätzers dienen soll, wurden ebenfalls Untersuchungen bezüglich der erwartbaren mittleren Akquisitionsperiode gemacht. Die von der Bildverarbeitung getrennten Untersuchungen, basierend auf Generierung synthetischer Kratererkennungs-Algorithmenergebnisse, lieferte dabei eine Basislinie, die daraufhin durch Versuche im TRON-Labor in einem realitätsnahen Umfeld mit Beleuchtung und Kamera qualifiziert werden konnten.

Das Verfahren erfüllt seinen Zweck, die Identifikation ausreichender Krater zur Poseakquisition, unter den folgenden breiten Rahmenbedingungen:

- Irreguläre Oberflächen
- Mittlere Kamera-Elevationen gegenüber der Oberfläche von  $5^\circ$  bis  $90^\circ$
- Mittlere Beleuchtungs-Elevation gegenüber der Oberfläche von  $5^\circ$  bis  $60^\circ$
- Für einen realistischen Trajektorien-Bilddatensatz erfolgreiche Akquisition im Mittel innerhalb von weniger als 3 verarbeiteten Bildern

Darüber hinaus erfüllen die Ergebnisse der Kraterdetektion alle Anforderungen, um unter Verwendung einer externen Poseschätzung als Daten für Prädiktor-Korrektor-Methoden der Absolutnavigation verwendet zu werden (wie im DLR-Projekt ATON) oder in Abwesenheit einer solchen Schätzung zur Reduzierung der Drift von Zustandsschätzern im Relativbetrieb über Methoden der visuellen Odometrie. Theoretisch könnten durch einen solchen Mehrfacheinsatz redundante Feature-Detektoren ersetzt werden, um so den auf einem Bordrechner laufenden Programmcode zu verschlanken.

Die Entwicklungslücke zur Einsetzbarkeit während einer echten Mondmission ist hauptsächlich in der Anpassungs- und Qualifizierungsarbeit der Softwareimplementierung zu sehen. Die für die Tests in dieser Dissertation verwendete Referenzimplementierung zeigt, dass die konzeptionell entwickelten Methoden durch Algorithmen abbildbar sind, die von ihrer Zeitkomplexität her geeignet sind zur Umsetzung auf raumfahrttauglicher Rechnerhardware.

## 6.2 Ausblick

Im Ausblick auf mögliche Weiterentwicklungen in der Zukunft sind zwei Hauptbereiche zu nennen: Zum einen eine Entwicklung in die Tiefe, also in Richtung einer Anwendbarmachung des hier vorgestellten Verfahrens für einen Missionseinsatz, zum anderen in die Breite, also eine Weiterentwicklung und Anpassung des Verfahrens für andere Einsatzbereiche.

Während einer Entwicklung in die Breite wären hauptsächlich andere Missionsszenarien zu untersuchen und die notwendigen Anpassungen in den Algorithmen vorzunehmen. Zu nennen sind dabei vor allem Missionen zu Großasteroiden. Die Herausforderungen hierbei sind die erheblich größere Unregelmäßigkeit der Zielkörper, es wird oft der gesamte Asteroid im Bild sein und nicht nur ein kleiner und relativ homogener Teil seiner Oberfläche, und die größere Unkenntnis über die Zielkörper, denn man muss davon ausgehen, dass keine so detaillierten Karten und Daten vorliegen, wie es mittlerweile für den Mond der Fall ist. Grundsätzlich sind die in dieser Arbeit entwickelten Verfahren aber auch für solche Fälle geeignet, denn sie wurden ja ausdrücklich mit minimalen Anforderungen an die beobachteten Oberflächen konzipiert. Auch sind Szenarien zur autonomen Kartierung denkbar, wobei mittels SLAM-Verfahren anhand erkannter Krater die im Vorhinein nur grob bekannten Kraterkarten vervollständigt werden können.

Die Entwicklung in die Tiefe auf der anderen Seite ist eher technischer als konzeptioneller Natur und gliedert sich in drei große Arbeitsbereiche:

**Methoden** — Während der Implementierung der Referenzsoftware blieben aus Zeitgründen überall kleinere Heuristiken oder Kniffe zur Abdeckung von Grenzfällen oder Entfernung von Outliern in Zwischendaten unimplementiert, die die Robustheit des Verfahrens weiter erhöhen können. In diesem Sinne ist die optimale Performance noch nicht erreicht. Im Rahmen statistischer Analysen wäre es weiterhin möglich, die Laufzeiten der Untermodule der Software zu untersuchen und durch Optimierung auf den nominellen Fall anstatt auf den Worst Case ihre mittlere Laufzeit um ein Vielfaches zu reduzieren.

**Implementierung** — Insbesondere die Auslagerung der bildverarbeitenden Algorithmen auf ein FPGA, algorithmische Restrukturierung der naiven Implementierung der in Abschnitt 3.9 vorgestellten Konstellations-Suchalgorithmen und allgemeine Anpassung hinsichtlich der Beschränkungen eines Flugrechnersystems (begrenzte Fließkomma-Rechenkapazität, harte Deadlines im Programmfluss, statische Speicherverwaltung) sind als Meilensteine auf dem Weg zum Flugeinsatz zu nennen.

---

**Qualifizierung** — Um die wahre Leistungsfähigkeit des Verfahrens und einer Implementierung zu ermitteln, ist letztendlich ein Test mit Bildmaterial notwendig, das tatsächlich aus dem Mondorbit während einer typischen Descent-Trajektorie aufgenommen wurde. Hier liegen leider bislang keine frei verfügbaren Datensätze in ausreichender Qualität und mit allen notwendigen Metainformationen (akkurate Absolutpose und innere Parameter der aufnehmenden Kamera) vor, anhand derer die Ergebnisse bezüglich der mittleren Akquisitionsperiode bestätigt werden könnten. Das TRON-Experiment, das für diese Arbeit durchgeführt wurde, kommt einem solchen Datensatz bislang am nächsten. In TRON ließen sich mit einigem Aufwand auch ein Teil der Posen von Kamera und Beleuchtungsquelle der systematischen Tests im Abschnitt 5.2 abbilden. Dies wäre interessant in der Hinsicht, dass durch die entstehenden Differenzen in den Erfolgsraten gegenüber den Tests mit durch Raytracing generierten Bildern ein Einblick in die Wirkung der Kameraeigenschaften und des Reflektionsverhaltens der Oberfläche möglich wären.





# Anhang A

## Beispielrechnungen

### A.1 Beispiel für eine lineare Nadirbestimmung

Es liege der Mondmittelpunkt in Kamerakoordinaten am Ort  $(0, -1, 2)^T$ , der Nadir ist also  $m = (0, -1/\sqrt{5}, 2/\sqrt{5})^T$ . Der Mondradius sei  $r = 1$ . Es seien drei Krater sichtbar, mit projektiven Richtungen und Normalen

$$\begin{aligned} c_1 &= (0, 0, 1)^T & \text{und } n_1 &= (0, 1, 0)^T, \\ c_2 &= \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T & \text{und } n_2 &= (0, 0, -1)^T, \\ c_3 &= \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)^T & \text{und } n_3 &= (1, 0, 0)^T, \end{aligned} \tag{A.1}$$

vergleiche Abbildung A.1: Die Matrix  $N$  ist dann

$$N = \begin{pmatrix} -1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \tag{A.2}$$

und der Eigenvektor zum betragskleinsten Eigenwert von  $N^T N$  ist gerade  $(0, -1/\sqrt{5}, 2/\sqrt{5})^T$ , also der Nadirvektor  $m$ .

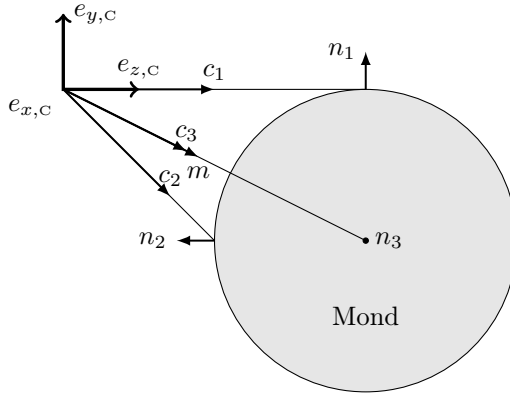


Abbildung A.1: Beispiel: Lineares Ausgleichsproblem für die Nadirrichtung

## A.2 Beispiel einer expliziten Oberflächenrekonstruktion in zwei Dimensionen

Es seien  $\mathbf{X}$  und  $\mathbf{Y}$  die Mengen der kartesischen Kamera- $x$ - und  $y$ -Koordinaten der in Abbildung 3.11 illustrierten planen Oberfläche für 5 Stützstellen:

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 2.5 & 3 & 3.5 & 4 \end{pmatrix} \quad (\text{A.3})$$

Die  $y$ -Achse ist hier die optische Achse des Systems. Die konfigurierten Zentren- und Normalenvektoren  $\mathbf{R}^{[\kappa]} = (\dots \{\hat{c}_i, \hat{n}_i\} \dots)$  sind damit

$$\begin{aligned} \mathbf{R}^{[\kappa]} = & \left( \left( \begin{pmatrix} 0.44721 \\ 0.89443 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right), \left( \begin{pmatrix} 0.37139 \\ 0.92848 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right), \right. \\ & \left( \begin{pmatrix} 0.31623 \\ 0.94868 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right), \left( \begin{pmatrix} 0.27472 \\ 0.96152 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right), \\ & \left. \left( \begin{pmatrix} 0.24254 \\ 0.97014 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right) \right) \quad (\text{A.4}) \end{aligned}$$

Wir interessieren uns hier nicht für die anderen möglichen Konfigurationen, der Anschaulichkeit halber wird schlicht die „wahre“ Konfiguration  $\kappa$  gewählt. Die zu den Werten (A.3) gehörenden Polarwinkel  $\Theta = (\dots \theta_i \dots)$  und Radien

$\mathbf{S} = (\dots r(\theta) \dots)$  sind dann in etwa

$$\begin{pmatrix} \boldsymbol{\Theta} \\ \mathbf{S} \end{pmatrix} = \begin{pmatrix} 1.1071 & 1.1903 & 1.2490 & 1.2925 & 1.3258 \\ 2.2361 & 2.6926 & 3.1623 & 3.6401 & 4.1231 \end{pmatrix}. \quad (\text{A.5})$$

Gleichung (3.53) ist in Polarkoordinaten für die Größe  $\frac{dr(\theta_i)}{d\theta_i} / r(\theta_i)$

$$\frac{dr(\theta_i)}{d\theta_i} / r(\theta_i) = \frac{\hat{n}_i^T}{\hat{c}_i^T \hat{n}_i} \begin{pmatrix} -\sin \theta_i \\ \cos \theta_i \end{pmatrix}, \quad (\text{A.6})$$

also an allen Stellen  $\boldsymbol{\Theta}$

$$\frac{dr(\boldsymbol{\Theta})}{d\boldsymbol{\Theta}} / r(\boldsymbol{\Theta}) = \begin{pmatrix} 2 & 2.5 & 3 & 3.5 & 4 \end{pmatrix}. \quad (\text{A.7})$$

In Abbildung A.2 sind die kartesischen Punkte  $\mathbf{X}$  gegen  $\mathbf{Y}$  aufgetragen, sowie an den Punkten jeweils die dortigen Normalen und Zentren-Richtungen  $\hat{n}_i$  und  $\hat{c}_i$  angezeichnet. In Abbildung A.3 sind die Stützstellen in Polarkoordinaten abgetragen.

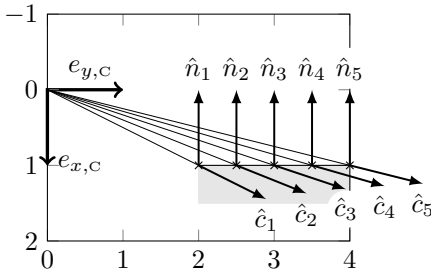


Abbildung A.2: kartesische Oberflächendaten mit Zentren- und Normalenvektoren

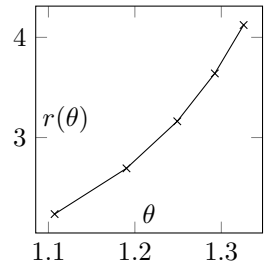


Abbildung A.3: Polare Oberflächendaten

Die „skalierte Ableitung“  $\frac{dr(\theta)}{d\theta} / r(\theta)$  nimmt an den Stützstellen die in Gleichung (A.7) angegebenen Werte an. Wird die rekonstruierte polare Oberflächenfunktion äquivalent zu Gleichung (3.54) definiert, benötigt man eine Methode, diese Daten zu integrieren. Dazu sei ein quadratischer Polynom-

satz gewählt:

$$A := \begin{pmatrix} \theta_1^2 & \theta_1 & 1 \\ \theta_2^2 & \theta_2 & 1 \\ \vdots & \vdots & \vdots \\ \theta_5^2 & \theta_5 & 1 \end{pmatrix} = \begin{pmatrix} 1.2258 & 1.1071 & 1 \\ 1.4168 & 1.1903 & 1 \\ 1.5601 & 1.2490 & 1 \\ 1.6705 & 1.2925 & 1 \\ 1.7578 & 1.3258 & 1 \end{pmatrix} \quad (\text{A.8})$$

Dann minimieren die Koeffizienten

$$F = \begin{pmatrix} 25.220 \\ -52.338 \\ 29.039 \end{pmatrix} \quad (\text{A.9})$$

an allen Stellen  $\Theta$  Quadrat von

$$A \cdot F - \begin{pmatrix} 2 & 2.5 & 3 & 3.5 & 4 \end{pmatrix}^\text{T} = \frac{dr(\Theta)}{d\Theta} \Big/ r(\Theta). \quad (\text{A.10})$$

Das Polynom  $25.220\theta^2 - 52.338\theta + 29.039$  interpoliert also nahezu die Daten aus der letzten Gleichung, sichtbar in Abbildung A.4. Eine Stammfunktion dieses Polynoms lässt sich durch „Aufleiten“ erhalten und Exponentiation liefert die rekonstruierte polare Oberflächenfunktion  $\hat{r}(\theta)$  (vgl. Gleichung (3.54))

$$\hat{r}(\theta) = \exp(8.40671\theta^3 - 26.16882\theta^2 + 29.03915\theta). \quad (\text{A.11})$$

Das Minimum dieser Funktion über alle Stützstellen  $\theta_i$  aus  $\Theta$  ist

$$\min_{\hat{c}_i} \{\hat{r}(\theta)\} = 96991.5146. \quad (\text{A.12})$$

Die Konstante  $\Lambda$  aus Gleichung (3.55) kann damit neutral als

$$\Lambda_{\text{ntr}} := \frac{1}{\min_{\hat{c}_i} \{r(\theta)\}} = 1.0310 \cdot 10^{-5} \quad (\text{A.13})$$

gewählt werden, oder, falls wie in diesem Beispiel  $\min_{\hat{c}_i} \{r(\theta)\} = 2.2361$  bekannt ist, kann die wahre Konstante  $\Lambda_{\text{opt}}$

$$\Lambda_{\text{opt}} := \frac{\min_{\hat{c}_i} \{r(\theta)\}}{\min_{\hat{c}_i} \{\hat{r}(\theta)\}} = 2.3405 \cdot 10^{-5} \quad (\text{A.14})$$

gewählt werden. Die resultierenden Rekonstruktionen der Oberfläche für diese beiden Konstanten in Polarkoordinaten sind in Abbildung A.5 gezeigt. Be-

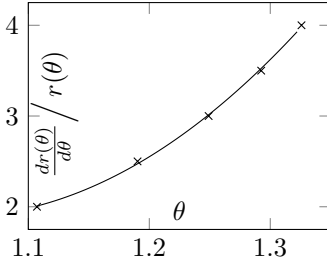


Abbildung A.4: Stützwerte der skalierten polaren Ableitung und ihr interpolierendes Polynom.

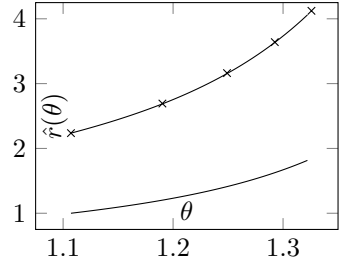


Abbildung A.5: Stützstellen und die integrierten Interpolierenden zu neutraler und optimaler Skala.

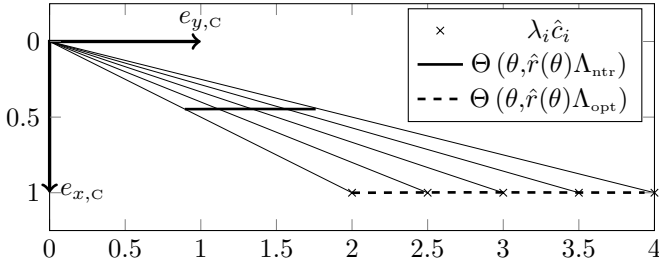


Abbildung A.6: Rekonstruierte Oberflächen für zu neutraler und optimaler Skala in kartesische Koordinaten zurücktransformiert.

zeichnet man mit  $\Theta(\theta, r) := (r \cos \theta, r \sin \theta)^T$  die Transformation von polaren nach kartesischen Koordinaten, dann entspricht die Rekonstruktion  $\hat{r}(\theta)\Lambda_{\text{ntr}}$  zur neutralen Konstante aus Abbildung A.5, zurücktransformiert in kartesische Koordinaten, qualitativ der Menge der Stützstellen  $\lambda_i \hat{c}_i$  (vgl. Abbildung A.6). Dies illustriert die in Gleichung 3.59 gemachte Aussage: Die Rekonstruktion der Oberfläche ist bis auf Skalierung korrekt. Die Rekonstruktion für den Skalar  $\Lambda_{\text{opt}}$  entspricht, wie erwartet, sogar der „tatsächlichen“ Oberfläche.

## A.3 Beispiel einer expliziten Spline-Oberflächenrekonstruktion in drei Dimensionen

Im obigen Beispiel A.2 wurde die Rekonstruierbarkeit und die Rekonstruktionstreue einer interpolierenden Funktion  $\hat{r}$  gezeigt, die gewählte Interpolante (das Polynom zweiten Grades) erfüllte dabei per Konstruktion schon die Minimale-Biegeenergie-Bedingung, die in den Abschnitten 3.7.7 und 3.7.8 (für das Problem höherer Ordnung) eingeführt wurden.

Am Beispiel eines minimalen Satzes von vier Kratern soll darüber hinaus jetzt für das Interpolationsproblem in drei Dimensionen die in Abschnitt 3.7 beschriebene Konstruktion des Interpolationssystems ausgeführt werden.

Als Beispieldatensatz seien dazu vier symmetrisch auf einer Ebene vor und unter der Kamera angeordnete Krater gegeben, die bei absoluten (kartesischen) Kamerakoordinaten  $(1.25, -1, 3.75)^T$ ,  $(1.25, -1, 6.25)^T$ ,  $(-1.25, -1, 3.75)^T$  und  $(-1.25, -1, 6.25)^T$  liegen, mit Normalen jeweils in Kamera- $y$ -Richtung (also gerade nach oben). Die Datentupel der Zentren- und Normalenvektoren  $\mathbf{r}_i = \{c_i, n_i\}$  und ihre Spiegelsymmetrischen  $\mathbf{r}'_i$  (vgl. Abschnitt 3.3.2, Unterpunkt „Anzahl der Lösungen für  $n$  und  $c$ “) sind dazu folglich:

$$\mathbf{r}_1 = \left( \begin{pmatrix} -1.9374 \cdot 10^{-1} \\ -1.5499 \cdot 10^{-1} \\ 9.6873 \cdot 10^{-1} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) \quad (\text{A.15})$$

$$\mathbf{r}'_1 = \left( \begin{pmatrix} -1.9374 \cdot 10^{-1} \\ -1.5499 \cdot 10^{-1} \\ 9.6873 \cdot 10^{-1} \end{pmatrix}, \begin{pmatrix} 6.0061 \cdot 10^{-2} \\ -9.5195 \cdot 10^{-1} \\ 3.0031 \cdot 10^{-1} \end{pmatrix} \right)$$

$$\mathbf{r}_2 = \left( \begin{pmatrix} -3.0651 \cdot 10^{-1} \\ -2.4521 \cdot 10^{-1} \\ 9.1974 \cdot 10^{-1} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) \quad (\text{A.16})$$

$$\mathbf{r}'_2 = \left( \begin{pmatrix} -3.0650 \cdot 10^{-1} \\ -2.4530 \cdot 10^{-1} \\ 9.1972 \cdot 10^{-1} \end{pmatrix}, \begin{pmatrix} 1.5039 \cdot 10^{-1} \\ -8.7968 \cdot 10^{-1} \\ -4.5116 \cdot 10^{-1} \end{pmatrix} \right)$$

$$\mathbf{r}_3 = \left( \begin{pmatrix} 3.0651 \cdot 10^{-1} \\ -2.4521 \cdot 10^{-1} \\ 9.1974 \cdot 10^{-1} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) \quad (\text{A.17})$$

$$\mathbf{r}'_3 = \left( \begin{pmatrix} 3.0650 \cdot 10^{-1} \\ -2.4530 \cdot 10^{-1} \\ 9.1972 \cdot 10^{-1} \end{pmatrix}, \begin{pmatrix} -1.5039 \cdot 10^{-1} \\ -8.7968 \cdot 10^{-1} \\ -4.5116 \cdot 10^{-1} \end{pmatrix} \right)$$

$$\begin{aligned}
 \mathbf{r}_4 &= \left( \begin{pmatrix} 1.9374 \cdot 10^{-1} \\ -1.5499 \cdot 10^{-1} \\ 9.6873 \cdot 10^{-1} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) \\
 \mathbf{r}'_4 &= \left( \begin{pmatrix} -1.9374 \cdot 10^{-1} \\ -1.5499 \cdot 10^{-1} \\ 9.6873 \cdot 10^{-1} \end{pmatrix}, \begin{pmatrix} -6.0061 \cdot 10^{-2} \\ -9.5195 \cdot 10^{-1} \\ -3.0031 \cdot 10^{-1} \end{pmatrix} \right)
 \end{aligned} \tag{A.18}$$

Sei aus den Daten (A.15) bis (A.18) die „nach oben“-Konfiguration

$$\mathbf{R}^{[\kappa]} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) \tag{A.19}$$

gewählt (vgl. Definition 3.1). Abbildung A.7 illustriert die räumliche Konfiguration der symmetrisch angeordneten vier Krater und ihre Projektion auf eine Bildebene. Die zu den Zentrendaten gehörenden virtuellen Bildkoordinaten

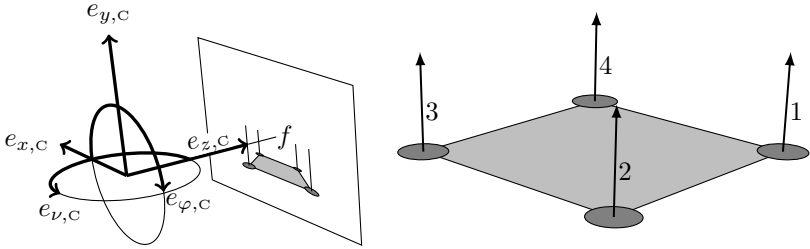


Abbildung A.7: Beispielszenario Spline-Oberflächenrekonstruktion räumlich

nach Definition 3.4 sind dann

$$\begin{aligned}
 \xi_1 &= (1.3758, -1.5866 \cdot 10^{-1})^T \\
 \xi_2 &= (1.2592, -2.6060 \cdot 10^{-1})^T \\
 \xi_3 &= (1.8824, -2.6060 \cdot 10^{-1})^T \\
 \xi_4 &= (1.7658, -1.5866 \cdot 10^{-1})^T
 \end{aligned} \tag{A.20}$$

Abbildung A.8 zeigt diese virtuellen Bildkoordinaten und eine resultierende Triangulation mit Kontrollnetz. Die Indizes des Vektors  $c^\Delta$  der einzigartigen Kontrollpunkte sind in Abbildung A.8 in den Kreisen zu finden. Zusätzlich sind dort jeweils die Kontrollpunkte  $c_\alpha^{t_{1,2}}$  der beiden einzelnen Dreiecke  $t_1$  und  $t_2$  notiert, zu denen die Indizes des globalen Kontrollvektors korrespondieren. Die Korrespondenz zu den linearen Kontrollpunkt-Indizes der beiden Einzeldreiecke kann in Tabelle 3.1 nachgeschlagen werden.

Die beiden  $10 \times 16$ -Permutationsmatrizen nach Gleichung (3.93) in Definition 3.10 sind damit

$$P_{t_1} = \begin{pmatrix} 0010000000000000 \\ 0000100000000000 \\ 0000010000000000 \\ 0100000000000000 \\ 0000001000000000 \\ 0000000100000000 \\ 0000000010000000 \\ 0000000001000000 \\ 0000000000100000 \\ 1000000000000000 \end{pmatrix}, \quad P_{t_2} = \begin{pmatrix} 0010000000000000 \\ 0000001000000000 \\ 0000000001000000 \\ 1000000000000000 \\ 0000000000010000 \\ 0000000000001000 \\ 0000000000000100 \\ 0000000000000010 \\ 0000000000000001 \\ 0001000000000000 \end{pmatrix}. \quad (\text{A.21})$$

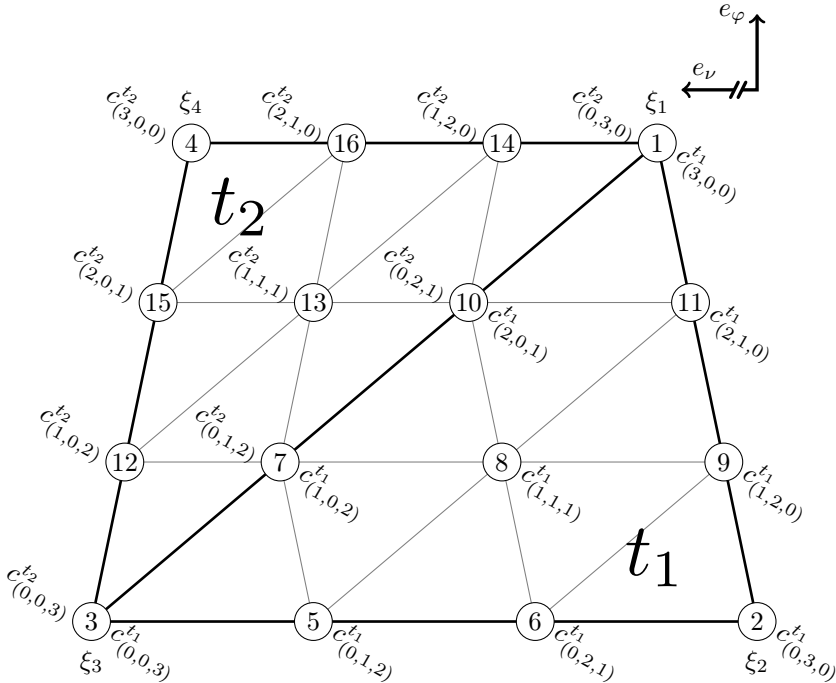


Abbildung A.8: Beispielszenario Spline-Oberflächenrekonstruktion virtuell

Zu diesen Daten mit der beschriebenen Triangulation  $\Delta = \{t_1, t_2\}$  werden jetzt die Übergangsbedingungen, die Interpolationsbedingungen und die Energie-Integralmatrix aufgestellt.



### Übergangsbedingungen $t_1$ nach $t_2$

Die vier  $\mathcal{C}^0$ -Übergangsbedingungen sind trivialerweise erfüllt, da bereits die redundanten Kontrollpunkte im Vektor  $c^\Delta$  zusammengefasst wurden. Die drei  $\mathcal{C}^1$ -Bedingungen lassen sich aus Tabelle 3.3 ablesen (Fall „L1-R2“, auf dem Kopf stehend zum dortigen Diagramm, da der Umlaufsinn der Dreiecks-Ecken hier andersherum ist):

$$\begin{aligned} c_{(1,0,2)}^{t_2} &= \left( c_{(1,0,2)}^{t_1}, c_{(0,1,2)}^{t_1}, c_{(0,0,3)}^{t_1} \right) \lambda^{t_1}(\xi_4) \\ c_{(1,1,1)}^{t_2} &= \left( c_{(2,0,1)}^{t_1}, c_{(1,1,1)}^{t_1}, c_{(1,0,2)}^{t_1} \right) \lambda^{t_1}(\xi_4) \\ c_{(1,2,0)}^{t_2} &= \left( c_{(3,0,0)}^{t_1}, c_{(2,1,0)}^{t_1}, c_{(2,0,1)}^{t_1} \right) \lambda^{t_1}(\xi_4) \end{aligned} \quad (\text{A.22})$$

Diese Gleichung ist äquivalent zu  $C^{\text{L1}} c^{t_2} = C^{\text{R2}} c^{t_1}$ , also mit  $\lambda^{t_1}(\xi_4) := (\lambda_1, \lambda_2, \lambda_3)^\top$  ausgeschrieben

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_{(0,0,3)}^{t_2} \\ c_{(0,1,2)}^{t_2} \\ c_{(0,2,1)}^{t_2} \\ c_{(0,3,0)}^{t_2} \\ c_{(1,0,2)}^{t_2} \\ c_{(1,1,1)}^{t_2} \\ c_{(1,2,0)}^{t_2} \\ c_{(2,0,1)}^{t_2} \\ c_{(2,1,0)}^{t_2} \\ c_{(3,0,0)}^{t_2} \end{pmatrix} = \begin{pmatrix} \lambda_3 & \lambda_2 & 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_3 & \lambda_2 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_3 & \lambda_2 & \lambda_1 \end{pmatrix} \begin{pmatrix} c_{(0,0,3)}^{t_1} \\ c_{(0,1,2)}^{t_1} \\ c_{(0,2,1)}^{t_1} \\ c_{(0,3,0)}^{t_1} \\ c_{(1,0,2)}^{t_1} \\ c_{(1,1,1)}^{t_1} \\ c_{(1,2,0)}^{t_1} \\ c_{(2,0,1)}^{t_1} \\ c_{(2,1,0)}^{t_1} \\ c_{(3,0,0)}^{t_1} \end{pmatrix} \quad (\text{A.23})$$

Mit den oben angegebenen Permutationen ergibt sich die Matrix  $S$  der  $\mathcal{C}^1$ -Bedingungen als

$$\begin{aligned} S &= C^{\text{R2}} P_{t_1} - C^{\text{L1}} P_{t_2} \\ &= \begin{pmatrix} 0 & 0 & \lambda_3 & 0 & \lambda_2 & 0 & \lambda_1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_3 & \lambda_2 & 0 & \lambda_1 & 0 & 0 & -1 & 0 & 0 & 0 \\ \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_3 & \lambda_2 & 0 & 0 & -1 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (\text{A.24})$$

$S$  ist nur drei Zeilen hoch, da es nur eine innere Kante in der Triangulation gibt.

### Interpolationsbedingungen

Die skalierten virtuellen Gradienten an den Ecken  $\xi_1$  bis  $\xi_4$  sind nach Gleichung (3.78)

$$\begin{aligned} G_1 &= (-1.975 \cdot 10^{-1}, 6.25)^T \\ G_2 &= (-3.2222 \cdot 10^{-1}, 3.75)^T \\ G_3 &= (3.2222 \cdot 10^{-1}, 3.75)^T \\ G_4 &= (1.975 \cdot 10^{-1}, 6.25)^T \end{aligned} \quad (\text{A.25})$$

Die  $\mathbf{B}_\xi$ -Matrizen lassen sich aus den Gleichungen (3.81), (3.82) und (3.83) ablesen. Die Gradienten an den Ecken  $\xi_1$  bis  $\xi_3$  werden bezüglich  $t_1$  ausgewertet,  $\xi_4$  bezüglich  $t_2$ , die Interpolationsgleichung ist damit

$$\begin{pmatrix} \mathbf{B}_\xi^{t_1}(1, 0, 0)P_{t_1} \\ \mathbf{B}_\xi^{t_1}(0, 1, 0)P_{t_1} \\ \mathbf{B}_\xi^{t_1}(0, 0, 1)P_{t_1} \\ \mathbf{B}_\xi^{t_2}(1, 0, 0)P_{t_2} \end{pmatrix} c^\Delta = \begin{pmatrix} G_1 \\ G_2 \\ G_3 \\ G_4 \end{pmatrix}. \quad (\text{A.26})$$

Die ausgeschriebene Gleichung ist in Abbildung A.10 dargestellt.

### Energiefunktionalmatrix

Es müssen die beiden Energiefunktionalmatrizen  $E^{t_1}$  und  $E^{t_2}$  gemäß Gleichung (3.92) für die beiden Dreiecke aufgestellt werden. Die resultierenden Matrizen sind leider im Format dieses Buches nicht mehr abbildbar. Das detaillierte Vorgehen ist aber das folgende:

1. Für die Dreiecke  $t_i$ ,  $i \in \{1, 2\}$ , mit ihren bestimmenden Koeffizienten  $a_{t_i}$  bis  $e_{t_i}$  aus Gleichung (3.69) müssen jeweils die  $3 \times 10$ -Matrizen der zweiten partiellen Ableitungen des Splines  $\mathbf{B}_{\xi\xi}^{t_i}(\boldsymbol{\lambda}(\xi))$  durch Einsetzen in Tabelle 3.5 gebildet werden.
2. Die Quadratur zur Lösung des Integrals wird wie in (3.92) gebildet, indem die drei Werte  $\boldsymbol{\lambda} = (0, 1/2, 1/2)$ ,  $(1/2, 0, 1/2)$ ,  $(1/2, 1/2, 0)$  in die  $\mathbf{B}_{\xi\xi}^{t_i}(\boldsymbol{\lambda})$  eingesetzt, die drei Matrizen addiert und durch ein Drittel der Fläche des betreffenden Dreiecks geteilt werden.
3. Die resultierenden Matrizen  $E^{t_i}$  werden mit den zum Dreieck gehörenden Permutationen  $P_{t_i}$  aus Gleichung (A.21) transponiert-links-rechts-multipliziert gemäß Gleichung (3.96) und beide Resultate werden zur System-Energimatrix  $E$  addiert.

Die in diesem Beispiel aufgestellten Matrizen  $E$ ,  $S$ ,  $I$ , und der Vektor  $G$  können jetzt benutzt werden, um das Minimierungsproblem (3.99) aufzustellen und zu lösen. Die eindeutige Lösung ist

$$c^\Delta = - \begin{pmatrix} 0 \\ 5.0151 \\ 6.3349 \\ 1.4293 \\ 7.0039 \\ 5.6844 \\ 5.6044 \\ 3.8100 \\ 3.8661 \\ 2.4574 \\ 2.0471 \\ 5.1857 \\ 3.5802 \\ 0.2567 \\ 3.4764 \\ 1.6860 \end{pmatrix} \cdot 10^{-1} \quad (\text{A.27})$$

In Abbildung A.9 ist der durch diesen Kontrollnetzvektor erzeugte Spline mit den Umrissen der Beispieloberfläche aus Abbildung A.7 dargestellt. Sichtbare Abweichungen rühren daher, dass nur vier Interpolationsstellen verwendet wurden. Die Interpolationstreue kubischer Polynome an große Flächen, die durch trigonometrische Funktionen dargestellt sind, ist in diesem Fall gering. Das hinzufügen weiterer Interpolationspunkte lässt solche Abweichungen schnell schrumpfen.

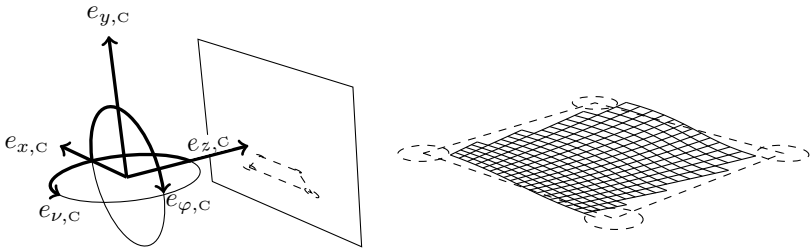


Abbildung A.9: Ergebnis der Spline-Oberflächenrekonstruktion im Beispiel

[illegible]

Abbildung A.10: Ausgeschriebene Interpolationsgleichung

# Literatur

- Ansar, Adnan und Yang Cheng (2005). „An analysis of spacecraft localization from descent image data for pinpoint landing on Mars and other cratered bodies“. In: *Photogrammetric Engineering & Remote Sensing* 71.10, S. 1197–1204.
- Ansar, Adnan und Kostas Daniilidis (2003). „Linear pose estimation from points or lines“. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 25.5, S. 578–589.
- Awanou, Gerard, Ming-Jun Lai und Paul Wenston (2005). „The Multivariate Spline Method for Scattered Data Fitting and Numerical Solutions of Partial Differential Equations“. In: *Wavelets and Splines*, S. 24–75.
- Besl, Paul J und Neil D McKay (1992). „Method for registration of 3-D shapes“. In: *Robotics-DL tentative*. International Society for Optics und Photonics, S. 586–606.
- Bredies, Kristian und Dirk Lorenz (2010). *Mathematische Bildverarbeitung: Einführung in Grundlagen und moderne Theorie*. Springer-Verlag.
- Cheng, Yang und Adnan Ansar (2005). „Landmark based position estimation for pinpoint landing on mars“. In: *Robotics and Automation, Proceedings of the 2005 IEEE International Conference on*. IEEE, S. 4470–4475.
- Cheng, Yang, Andrew E Johnson u. a. (2003). „Optical landmark detection for spacecraft navigation“. In: *Advances in the Astronautical Sciences* 114.3, S. 1785–1803.
- Cheng, Yang und JK Miller (2003). „Autonomous landmark based spacecraft navigation system“. In: *Advances in the Astronautical Sciences* 114.3, S. 1769–1783.
- Cobham Gaisler AB (o.D.). *Produktseite „GR712RC Dual-Core LEON3FT SPARC V8 Processor“*: Abgerufen am 30. Nov. 2015. URL: <http://www.gaisler.com/index.php/products/components/gr712rc>.

- Cross, AM (1988). „Detection of circular geological features using the Hough transform“. In: *International Journal of Remote Sensing* 9.9, S. 1519–1528.
- Delaune, Jeff, Guy Le Besnerais, Martial Sanfourche, Aurélien Plyer u. a. (2011). „Camera-aided inertial navigation for pinpoint planetary landing on rugged terrains“. In: *International Planetary Probe Workshop*.
- Delaune, Jeff, Guy Le Besnerais, Martial Sanfourche, Thomas Voirin u. a. (2012). „Optical terrain navigation for pinpoint landing: Image scale and position-guided landmark matching“. In: *Advances in the Astronautical Sciences* 144, S. 627–643.
- Farin, Gerald E (1979). „Subsplines über Dreiecken“. Dissertation. Universität Braunschweig.
- (1986). „Triangular Bernstein-Bézier patches“. In: *Computer Aided Geometric Design* 3, S. 83–127.
- Forsyth, David u. a. (1991). „Invariant descriptors for 3D object recognition and pose“. In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 10, S. 971–991.
- Gao, Xiao-Shan u. a. (2003). „Complete solution classification for the perspective-three-point problem“. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 25.8, S. 930–943.
- Gold, Steven u. a. (1998). „New algorithms for 2D and 3D point matching: Pose estimation and correspondence“. In: *Pattern recognition* 31.8, S. 1019–1031.
- Hale, Matthew J (2007). „Terrain relative localization for lunar entry, descent, and landing“. Master Thesis. Massachusetts Institute of Technology.
- Hanak, Francis Chad (2009). „Lost in low lunar orbit crater pattern detection and identification“. Dissertation. The University of Texas at Austin.
- Hanak, Francis Chad, Tim Crain und Robert Bishop (2010). „Crater Identification Algorithm for the Lost in Low Lunar Orbit Scenario“. In: *Advances in the Astronautical Sciences* 137.32.
- Hartley, Richard und Andrew Zisserman (2003). *Multiple view geometry in computer vision*. Cambridge university press.
- Higham, Nicholas J (2002). *Accuracy and Stability of Numerical Algorithms*. 2. Auflage. SIAM.
- Kato, Manabu, Susumu Sasaki, Yoshisada Takizawa u. a. (2010). „The Kaguya mission overview“. In: *Space science reviews* 154.1-4, S. 3–19.
- Krüger, Hans u. a. (2014). „On-Ground Testing Optical Navigation Systems for Exploration Missions“. In: *9th International ESA Conference on Guidance, Navigation & Control Systems*.

- Kudari, Vishwas (2015). „Development and Implementation of Methods for Mapping Lunar Impact Craters by Optical Means“. Masterarbeit. Hochschule Darmstadt.
- Lepetit, Vincent, Francesc Moreno-Noguer und Pascal Fua (2009). „EPnP: An accurate  $\mathcal{O}(n)$  solution to the PnP problem“. In: *International journal of computer vision* 81.2, S. 155–166.
- Leroy, Bertrand u. a. (2001). „Crater detection for autonomous landing on asteroids“. In: *Image and Vision Computing* 19.11, S. 787–792.
- Maass, Bolko (2010). „Entwicklung mathematischer Methoden der Bildverarbeitung zur optischen Erkennung von Mondkratern“. Diplomarbeit. Universität Bremen.
- Maass, Bolko, Hans Krüger und Stephan Theil (2011). „An edge-free, scale-, pose- and illumination-invariant approach to crater detection for spacecraft navigation“. In: *Image and Signal Processing and Analysis (ISPA), 2011 7th International Symposium on*. IEEE, S. 603–608.
- Mammarella, Marco u. a. (2011). „Advanced Optical Terrain Absolute Navigation for Pinpoint Lunar Landing“. In: *Advances in Aerospace Guidance, Navigation and Control*. Springer, S. 419–430.
- Matas, Jiri u. a. (2004). „Robust wide-baseline stereo from maximally stable extremal regions“. In: *Image and vision computing* 22.10, S. 761–767.
- Neukum, Gerhard (1977). „Different ages of lunar light plains“. In: *The Moon* 17.4, S. 383–393.
- Neukum, Gerhard, Beate König und Jafar Arkani-Hamed (1975). „A study of lunar impact crater size-distributions“. In: *The Moon* 12.2, S. 201–229.
- opencv.org (o.D.). *OpenCV 2.4.11.0 documentation*. Abgerufen am 4. Aug. 2015. URL: [http://docs.opencv.org/modules/calib3d/doc/camera\\_calibration\\_and\\_3d\\_reconstruction.html#solvePnP](http://docs.opencv.org/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html#solvePnP).
- Paproth, Carsten u. a. (2012). „SENSOR++: simulation of remote sensing systems from visible to thermal infrared“. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXII ISPRS Congress*, S. 257–260.
- Penate-Sanchez, Adrian, Juan Andrade-Cetto und Francesc Moreno-Noguer (2013). „Exhaustive linearization for robust camera pose and focal length estimation“. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35.10, S. 2387–2400.
- Rangarajan, Anand, Haili Chui und Fred L Bookstein (1997). „The softassign procrustes matching algorithm“. In: *Information Processing in Medical Imaging*. Springer, S. 29–42.

- Samaan, Malak A, Daniele Mortari und John L Junkins (2006). „Nondimensional star identification for uncalibrated star cameras“. In: *The Journal of the Astronautical Sciences* 54.1, S. 95–111.
- Shoemaker, E. M. u. a. (1970). „Geology of the Apollo 12 Landing Site“. In: *Apollo 12 Preliminary Science Report NASA SP-235*, S. 113–157.
- Simard Bilodeau, Vincent (2015). „Autonomous Vision-Based Terrain-Relative Navigation for Planetary Exploration“. Dissertation. University of Sherbrooke.
- Simard Bilodeau, Vincent, Sébastien Clerc, Jean De Lafontaine u. a. (2010). „A Vision-Based Navigation Algorithm for Pin-Point Landing“. In: *Advances in the Astronautical Sciences* 137.37.
- Simard Bilodeau, Vincent, Sébastien Clerc, Rémi Draï u. a. (2014). „Optical Navigation System for Pin-Point Lunar Landing“. In: *The International Federation of Automatic Control, the World Congress of.* Bd. 19. 1, S. 10535–10542.
- Simard Bilodeau, V u. a. (2012). „Pinpoint lunar landing navigation using crater detection and matching: Design and laboratory validation“. In: *Proceedings of the AIAA Guidance, Navigation, and Control Conference.*
- Singh, Leena und Sungyung Lim (2008). „On lunar on-orbit vision-based navigation: terrain mapping, feature tracking driven EKF“. In: *Proceedings of the AIAA Guidance, Navigation and Control Conference.*
- Spigai, Marc, Sébastien Clerc und Vincent Simard Bilodeau (2010). „An image segmentation-based crater detection and identification algorithm for planetary navigation“. In: *Intelligent Autonomous Vehicles, Lecce, Italy.*
- Strobl, Klaus u. a. (o.D.). *DLR CalDe und DLR CalLab*. Institut für Robotik und Mechatronik, DLR. URL: <http://www.robotic.dlr.de/callab/>.
- Theobald, Douglas L (2005). „Rapid calculation of RMSDs using a quaternion-based characteristic polynomial“. In: *Acta Crystallographica Section A: Foundations of Crystallography* 61.4, S. 478–480.
- Van Pham, Bach, Simon Lacroix und Michel Devy (2008). „Landmarks Constellation Based Position Estimation for Spacecraft Pinpoint Landing“. In: *Advanced Space Technologies in Robotics and Automation (ASTRA), 10th Symposium on.*
- Van Pham, Bach, Simon Lacroix, Michel Devy u. a. (2010). „Fusion of absolute vision-based localization and visual odometry for spacecraft pinpoint landing“. In: *Proceedings of the 61st International Astronautical Congress.* International Astronautical Federation.
- Varma, Manik und Andrew Zisserman (2004). „Estimating illumination direction from textured images“. In: *Proceedings of the 2004 IEEE Computer*



- Society Conference on Computer Vision and Pattern Recognition*. Bd. 1. IEEE, S. I–179.
- Wertz, James R (1978). *Spacecraft attitude determination and control*. Bd. 73. Kluwer Academic Publishers.
- Wokes, D S und S M Wokes (2010). „Surveying and Pose Estimation of a Lander Using Approximative Crater Modelling“. In:
- Wood, CA und L Anderson (1978). „New morphometric data for fresh lunar craters“. In: *Lunar and Planetary Science Conference Proceedings*. Bd. 9, S. 3669–3689.
- Yu, Meng, Hutao Cui und Yang Tian (2014). „A new approach based on crater detection and matching for visual navigation in planetary landing“. In: *Advances in Space Research* 53.12, S. 1810–1821.
- Zhang, Linbo, Tao Cui und Hui Liu (2009). „A Set of Symmetric Quadrature Rules on Triangles and Tetrahedra“. In: *Journal of Computational Mathematics* 27.1, S. 89–96.